Behavioral estimates of human frequency selectivity at low frequencies

Orellana, Carlos Andrés Jurado

Publication date:
2011

Document Version
Early version, also known as pre-print

Behavioral estimates of human frequency selectivity at low frequencies

Ph.D. thesis by
Carlos Jurado

Section of Acoustics
Department of Electronic Systems, Aalborg University
Denmark

October 2010

Behavioral estimates of human frequency selectivity at low frequencies

ISBN: 978-87-92328-69-4

 Address correspondence to:
Section of Acoustics, Department of Electronic Systems
Aalborg University, Frederik Bajers vej 7-B5, DK-9220 Aalborg
Phone: +45 9940 8710, E-mail: acoustics@es.aau.dk

Ph.D. thesis defended on April 28, 2011
Ph.D. degree awarded on June 8, 2011
Aalborg University, Denmark

Assessment Committee:
Associate Professor Flemming Christensen, Aalborg University, Denmark (chairman)
Professor Hugo Fastl, Technische Universität München, Germany
Professor Yôiti Suzuki, Tohoku University, Japan

Supervisors:
Professor Henrik Møller, Aalborg University, Denmark
Assistant Professor Christian S. Pedersen, Aalborg University, Denmark

Preface

This thesis is submitted to the Faculty of Engineering, Science and Medicine at Aalborg University, Denmark, in partial fulfillment of the requisits for the PhD degree. The work described here was carried out at the Section of Acoustics, Aalborg University, between September 2007 and September 2010. The PhD project was part of the framework program 26-00-0095: “Research in human sound perception – with special reference to electro acoustic applications” under the Danish Research Council for Technology and Production (FTP).

I would like to express my gratitude to my supervisors Henrik Møller and Christian Sejer Pedersen for their valuable advice, share of experience, and impecable scientific criterion. I'm additionally grateful to Christian for his translation of the summary into Danish. I am also deeply thankful to Brian C. J. Moore for his large enthusiasm with the project that resulted in a fruitful collaboration and learning experience. I also give thanks to Torsten Marquardt for his spontaneous involvement in the project that also led to a worthy cooperation.

I appreciate the vital assistance of Claus Vestergaard and Peter Dissing in the laboratory; the indispensable help of Mette Billeskov, Lone Engen, and Linda Villadsen for their efficient organization of accounting and administrative matters; and the overall positive support of our section head, Dorte Hammershøi, to the project and its finantial needs.

Finally, I thank all colleagues that helped me in one way or another and the motivated participants of the listening tests.

Carlos A. Jurado
October 2010
Aalborg, Denmark

Table of contents

Summary

A fundamental property of the auditory system is its frequency resolving power. This allows us to process sound in such a way as to provide an effective frequency analysis of it. Understanding the properties of specific filtering units, the *auditory filters*, across frequency, has been essential in the development of auditory models that describe our perception of sound. Unfortunately, little has been known about the characteristics of frequency selectivity in the low-frequency range (i.e. below 200 Hz) and practically no data existed for frequencies below 100 Hz, due to various complicating factors. Nevertheless, a proper description of frequency selectivity at low frequencies has long time been needed, not only to advance our limited knowledge, but in light of the many problems produced by low-frequency noise.

The subject of this PhD thesis concerns a detailed characterization of human frequency selectivity in the low-frequency range. A series of experiments have been carried out with this aim. Careful considerations were necessary to provide a proper control of the acoustic signals used in the experiments. Modifications were also necessary in aspects of existing methodologies for their applicability in the low-frequency range. As well, special attention has been given to factors thought to influence low-frequency hearing, such as the filtering effects attributed to the middle-ear-transfer function (METF).

In the first experiment, characteristics of frequency selectivity were compared across frequency, considering signal frequencies (f_s) from 50 up to 1000 Hz, using the notched-noise method. To obtain an adequate description of the lower filter skirt at the lowest values of f_s, the lower flanking band had to be emphasized. A main outcome from this experiment was to infer the high degree of influence the METF has on tuning (the latter transfer function was assumed to include the effects of the helicotrema shunt). Results suggested that the METF increasingly sharpens (i.e. defines) the lower skirt of the auditory filter, especially in the frequency range below 100 Hz. Also, in this range, the efficiency in the detection process was found to moderately improve.

A second experiment was carried out to extend results down to f_s = 31.5 Hz, while also providing a setup that allowed higher masker levels without distortion. The method was to measure a psychophysical tuning curve, which could provide more direct estimates of the shape of the auditory filter. This time, rough estimates of the METF were obtained for each subject by measuring an equal-loudness contour (ELC). The results were in overall agreement with those of the previous experiment. The use of lower values of f_s together with the ELCs allowed to resolve the center frequency (CF) of the most apical auditory filter, which appears to be located between about 40 to 50 Hz. Signals below that would be detected via the low-frequency skirt of this “bottom” auditory filter.

In both experiments it was found that tuning was affected for f_s below 80 Hz, which made the bandwidth of the auditory filter increase below that. An analysis of the possible effect on tuning of individual METFs (using objective estimates of the latter) could largely explain this phenomenon. Furthermore, individual differences could be expected to increase with decreasing f_s, as was observed in the psychoacoustical measures.

In a third experiment the relationship between perceived loudness for sinusoids and objective estimates of the METF obtained from distortion-product-isomodulation curves (DPIMC) was investigated. The outcome suggested that these are closely connected, and that the specific frequency dependence of the ELCs is not accounted for in standardized isophon curves. Although qualitatively similar, the ELCs were steeper than the DPIMCs, especially below about 40 Hz. This comparison allowed to improve the interpretation of the auditory filtering process inferred from the previous experimental results.

The evidence found in this work suggests that the helicotrema increasingly affects auditory tuning as CFs approach the apical end of the cochlea; decreasing the tuning power of the hearing organ and setting a limit as to where the auditory filter with the lowest CF can be located. Results are expected to contribute to the further development of auditory models –and standardized isophon curves, for their proper applicability in the low-frequency range.

Resumé (Summary in Danish)

En fundamental egenskab ved de auditoriske system er dets evne til at adskille frekvenser. Dette gør os i stand til at processere lyd på en måde, som giver en effektiv frekvensanalyse af lyden. Det har været essentielt at forstå egenskaberne af specifikke filtreringsenheder, auditorisk filter ved forskellige frekvenser, for at udvikle modeller af det auditoriske system, der beskriver vores lydopfattelse. Desværre har der været begrænset viden om karakteristika af frekvensselektivitet, i det lavfrekvente frekvensområde (under 200 Hz) and der eksisterer praktisk taget ingen data for frekvenser under 100 Hz, pga. forskellige komplicerende faktorer. En fyldestgørende af frekvensselektivitet ved lave frekvenser har længe været nødvendig - ikke kun for at udvide vores begrænsede viden - men set i lyset af de mange problemer, som lavfrekvent støj medfører.

Denne PhD-afhandling indeholder en detaljeret karakterisering af menneskets frekvensselektivitet i det lavfrekvente område. En række eksperimenter er blevet udført med dette mål for øje. Omhyggelige overvejelser var nødvendige for at have fuldt kontrol over de akustiske signaler, der blev brugt i eksperimenterne. Modifikationer af eksisterende metodologier var også nødvendige for at anvende disse i det lavfrekvente område. Desuden er særlig opmærksomhed givet til faktorer, der menes at have indflydelse på hørelsen ved lave frekvenser, såsom filtrerings effektor, som tillægges mellemørets overføringsfunktion (METF).

I det første eksperiment blev frekvensselektivitet sammenlignet på tværs af frekvens med signalfrekvens (f_s) fra 50 til 1000 Hz, ved brug af notched-noise metoden. For at opnå en fyldestgørende beskrivelse af de nedre filter skørter for de laveste værdier af f_s var det nødvendigt at lægge større vægt på det nedre flankerende bånd. Et hovedresultat fra dette eksperiment var at udlede inflydelsen som METF har på tuning (overføringsfunktion antagedes at inkludere kortslutningseffekten af helicotrema). Resultaterne viser, at METF bliver skarpere og skarpere (dvs. definerer) det lave skørt af det auditoriske filter, især i frekvensområdet under 100 Hz. I dette område sås en moderat forbedring i effektiviteten af detektionsprocessen.

I det andet eksperiment blev frekvensområdet udvidet ned til f_s = 31.5 Hz, i en forsøgsopstilling, som tillader højere maskeringsniveauer uden forvrængning. Metoden var psykofysiske tuning kurver, hvorfra et mere direkte estimate af udformningen af de auditoriske filtre kunne uddrages. Denne gang blev simple estimater af METF fundet for hver forsøgsperson ved at måle en hørestyrkekurve (ELC). De overordnede resultater var de samme som i det forrige eksperiment. Brugen af lavere værdier for f_s sammen med ELC gjorde det muligt at finde center frekvensen (CF) for det mest apikale
auditoriske filter, som viser sig at ligge mellem 40 og 50 Hz. Signal under dette frekvensområde vil blive detekteret via det lavfrekvente skørt af dette laveste filter.

I begge eksperimenter sås det, at tuning var påvirket for f_s under 80 Hz, så båndbredden af det auditoriske filter øges under dette. En analyse af mulige effekter på tuning fra individuelle METFs (ved brug af objektive estimater af disse) kunne i stor grad forklare

dette fænomen. Desuden kunne det forventes at individuelle forskelle ville øges ved lavere f_s , hvilket også kunne ses i de psykoakustiske estimater.

I et tredje eksperiment blev sammenhængen mellem opfattet hørestyrke for toner og objektive estimater af MEFT fra distortion-product-isomodulation curves (DPIMC) undersøgt. Udfaldet viser, at disse er tæt forbundet, og at den specifikke frekvensafhængihed af ELCs ikke er afspejlet i de standardiserede isophon kurver. Selvom de var kvalitativt ens, så var ELCs stejlere end DPIMCs, især under ca. 40 Hz. denne sammenligning hjælper til at forbedre fortolkningen af de auditoriske filtreringsprocess, fra de tidligere eksperimenters data.

Evidens fundet i disse undersøgelser indikerer at helicotrema i stigende omfang påviker auditorisk tuning når CFs nærmer sig den apikale ende af cochlea, reducerer tunings effekten af høremekanismen og sætter en grænse for hvor det auditoriske filter med den laveste CF kan lokaliseres. Resultaterne forventes at bidrage til yderligere udvikling af auditoriske modeller – og standardiserede isophon kurver, for deres behørige brugbarhed i det lavfrekvente område.

Introduction

Frequency selectivity refers to the ability our hearing organ has to resolve frequency components of complex sounds. This capacity allows us to, for example, break down complex sound such as speech or music into different spectral components, making it possible to perceive –and make use of– the richness of natural sound phenomena. Frequency resolving power is a fundamental attribute of the auditory system, which has evolved not only in humans and other mammals, but that is present too in other vertebrates (Manley, 2000). A common way to conceptualize the frequency-selective behavior of the ear is to represent the peripheral auditory system as it was composed by a set of overlapping bandpass filters, the *auditory filters*, whose bandwidth and shape are frequency, and somewhat level, dependent.

Auditory models, aiming to characterize human perception of sound, make extensive use of information about frequency selectivity (see e.g. Moore *et al.*, 1997). However, few studies existed that described the characteristics of the auditory filter in the low-frequency range, i.e. below 200 Hz. Furthermore, below 100 Hz, results were practically inexistent. This is perhaps explained in part by practical difficulties that arise when working with low-frequency acoustic signals, or, possibly also, due to complication in the treatment and interpretation of factors influencing low-frequency hearing. However, our limited knowledge is critical if we consider that many environmental sounds contain considerable low-frequency energy. Such sounds, due to their large wavelenghts, can easily propagate over long distances and produce problems with annoyance and masking (Leventhall, 2004).

From this perspective, a characterization of the auditory filter shape at low frequencies and of other key factors that determine our perception of sound in the low-frequency range has been considered useful and necessary. Such descriptions are thought to be fundamental in order to understand, model, and potentially predict important perceptual attributes, such as the audibility, loudness, and masking produced by complex low-frequency sounds.

In this PhD thesis, a detailed description of auditory frequency selectivity with focus in the low-frequency range is given. The main aims have been to answer the following questions:

(1) What is the shape and bandwidth of the auditory filter and how do they vary with frequency in the low-frequency range?

(2) What are the factors that affect frequency selectivity at low frequencies and how do they affect it?

(3) What is the degree of individual differences in frequency selectivity at low frequencies and which factors may influence them?

(4) Is there a limiting low frequency where an auditory filter is or can be centered?

An additional and related aim has been to establish in detail how perceived loudness for pure low-frequency tones varies with frequency and to evaluate its possible relation to the shape of the middle-ear-transfer function. The latter transfer function includes a cochlear component at low frequencies (see e.g. Marquardt and Hensel, 2008), and such comparison has therefore been important for the interpretation of the results describing frequency selectivity.

In the the following section, a further description of the subject of frequency selectivity and its relationship to masking is given. Following this, sub-section 1.1 presents a description of the concepts of *critical band* and *auditory filter* together with a brief background of previous work. Later on, sub-section 1.2 describes further general considerations to take into account regarding frequency selectivity and the perception of complex sounds. In section 2, relevant methods useful to characterize frequency selectivity are described. General challenges associated with the low-frequency are discussed in sub-section 2.6.

1 Frequency selectivity and masking

Our perception of a sound signal does not solely depend on its own frequency and intensity, but also on the presence of other sounds. A common experience is that we cannot distinguish or “hear out” particular sounds because we hear, instead, other sounds that are present. In other words, the threshold of audibility of a particular sound may be raised due to the presence of another, “masking”, sound. This phenomena is known simply as *masking*. A well known property of masking is that a sound is easily masked by another if they are relatively close in frequency, while the amount of masking tends to decrease as their frequencies differ. Following this and other similar observations, it has been suggested that masking may be a direct consequence of how the the acoustic signals are internally represented in the hearing organ (i.e. how they are coded), which appears to be at least partially dependent on the frequency resolving power of the basilar membrane (BM). It should be noted, however, that it is not fully understood how exactly and how far in the auditory system the internal mechanisms of frequency selectivity and/or masking operate. What has been previously assumed, based on a diversity of evidence (see examples below), is that when masking occurs, this would happen because the limits of the frequency selectivity of the hearing organ are reached.

Although a variety of psychoacoustic phenomena are associated with frequency selectivity (see e.g. Zwicker *et al.*, 1957 ; Moore, 2003), masking is usually where it is most directly demonstrated and measured. A strong indication that measures based on masking can provide a good approximation of the frequency selectivity of the auditory system is that many similarities exist between the frequency selectivity measured in the BM and psychophysical estimates of frequency selectivity derived from masking experiments (Moore, 1986). For example, (physiological) neural-tuning curves present very similar characteristics as observed in estimates of auditory tuning derived psychophysically (Pickles, 1984; Palmer, 1987). As well, measures of the bandwidth of the auditory filter, when estimated both behaviorally and neurophysiologically on the same species, have been found to be similar (Evans *et al.*, 1989). These results have

suggested that the auditory filters have their origins in the periphery, and that the BM provides at least an initial basis for the filtering process leading to frequency selective behavior.

Regarding the mechanism of masking, there are two notions of how it may operate, namely *swamping* and *suppression* (Moore, 2003). In the first one, masking would occurr when the neural activity elicitated by the masker is high enough in the channels (i.e. auditory filters) where the signal is normally detected, and therefore the relatively lower activity produced by the signal would render it undetectable. In other words, the masker overwhelms or *swamps* these channels, so that the activity added by the signal is insignificant. On the other hand, *suppression* involves the cease of activity in the channels where the signal is normally detected due to the simultaneous presence of the masker. This has been found to occur for masker frequencies well below and above the signal frequency (see e.g. Sachs and Kiang, 1968; Arthur *et al.*, 1971; Ruggero *et al.*, 1992).

However, it is generally thought that masking reflects rather a combination of these mechanisms than any of these in isolation (Moore and Vickers, 1997; Gifford and Bacon, 2000).

Measures of frequency selectivity derived from non-simultaneous masking experiments, such as forward masking, are thought not to be affected or to be less affected by suppression effects than measures derived from simultaneous masking experiments (Delgutte, 1990). Forward masking can be, however, much susceptible to spectral splatter effects (Leshowitz and Wightman, 1972; Moore, 1981) produced from gating the signals and maskers. This may be critical at low frequencies, due to the generally longer ramps (and signals) required, and the fact that in absolute terms the frequency selectivity of the ear is better than at higher frequencies –and therefore probably more sensitive to splatter effects across channels. In addition, due to the longer ramps required, confusion may occur between signals and maskers (Terry and Moore, 1977; Moore, 1981; Lopez-Poveda *et al.*, 2003).

Because of suppression effects, measures of frequency selectivity derived from simultaneous masking may underestimate the underlying frequency selectivity of the hearing organ (Heinz *et al.*, 2002; Moore, 2002; Shera *et al.*, 2002). This is because the signal may be supressed by the presence of masker, making the latter more effective in masking the signal than otherwise. In spite of this, however, simultaneous masking is thought to provide “effective” measures of frequency selectivity, where non-linear effects such as suppression, which are probably present too in the mayority of cases, are included. Further, these estimates have proven to be a succesful tool in auditory models that can predict the audibility and loudness of both sinusoidal and complex sounds (see e.g. Moore *et al.*, 1997; ANSI, 2007).

1.1 The concepts of critical bandwidth and auditory filter

The concept of *critical bandwidth* was first introduced by Fletcher (1940). He performed a "band-widening" experiment, in which masked threshold for a pure tone centered in a band of noise of constant power density N_0 (the masker) is measured as a function of the bandwidth of the noise. The total power of the noise thus increased with increasing bandwidth. Fletcher found that the power of the tone at masked threshold, P_s, increased proportional to the bandwidth –and therefore power– of the noise, up to a critical point, after which further increases in noise bandwidth did not affect threshold anymore. He called this breakpoint the *critical bandwidth* (CB) and found that it increased as he increased the center frequency of the noise band (with the tone always centered in the noise band). The observation is summarized in the following equation:

$$P_s = K \cdot CB \cdot N_0 \qquad (1),$$

where K is a proportionality constant indicating the required S/N to achieve detection. Based on this evidence, and following Helmholtz (1863), Fletcher suggested that the peripheral auditory system behaves as it was composed by a set of overlaping bandpass filters, the *auditory filters*, with the CB being a measure of their effective bandwidth. Figure 1 illustrates the basic procedure used in the band-widening experiment described above.

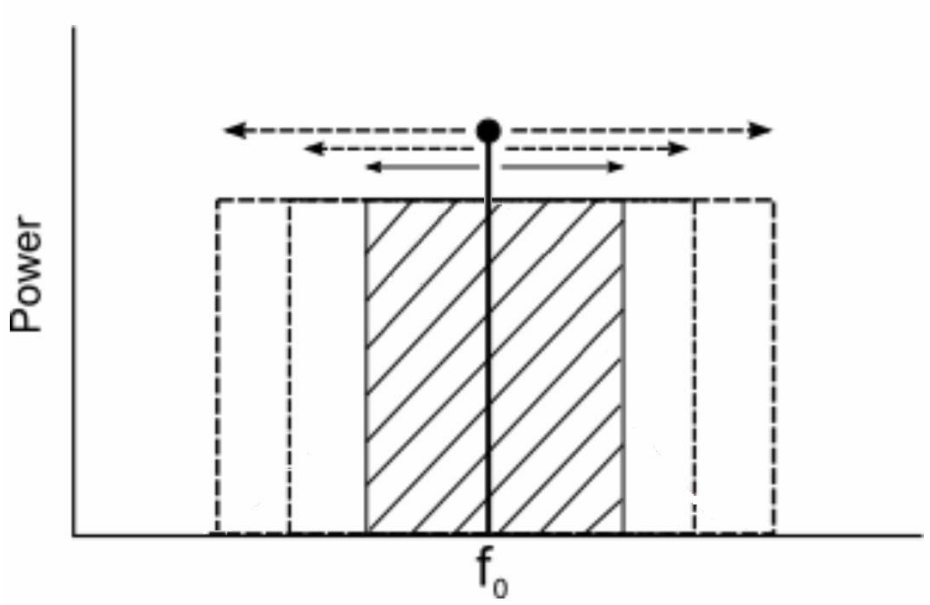

Figure 1: Spectral configuration of the signal (a tone with frequency f_0, shown in vertical-solid line) and masker (shaded-area box) used in a band-widening experiment. The arrows indicate increases in masker bandwidth.

In the model of masking assumed by Fletcher it is presupposed that: a) only spectral components around the tone signal, falling within a CB, affect its detection, while components outside the CB are highly attenuated and thus do not influence detection; and b) when the masker just masks the signal, the ratio of the powers of the tone and masker

is a constant, *K*. This model has been later called the "power-spectrum model of masking" (Patterson and Moore, 1986).

Incorrect assumptions regarding the value of *K* across frequency and scarce data prevented Fletcher in doing accurate CB estimations. However, the principles of his experimental approach and practical approximations used to describe the frequency selective behavior of the auditory system have proven to be major contributions. The ratio $P_s/N_0 = K \cdot CB$ was later called the *critical ratio*, and subsequently several studies estimated the CB from this ratio, assuming wrongly that K = 1 (i.e. 0 dB) for all frequencies (Hawkins and Stevens, 1950; Bilger and Hirsh, 1956; Green *et al.*, 1959). Although such assumption made measurements faster (no band widening was required), more recent data suggests that *K* is typically equal to 0.4 at mid frequencies, and does vary with center frequency, typically increasing at low frequencies.

Similar band-widening experiments as performed by Fletcher have been done subsequently (Schaeffer *et al.*, 1950; Hamilton, 1957; Bos and de Boer, 1966; Spiegel, 1981). Although they have included some variations in experimental procedure, they all are in accordance with the pattern of results described above, i.e. that of an increase in P_s as the noise bandwidth is increased, up to a critical point, the CB. However, evidence of not-so-abrupt transitions (i.e. breackpoints where threshold flattens) was found, indicating that the rough approximation of a rectangular filter shape assumed in the CB model was unrealistic and could be improved. This is how the *auditory filter* concept was born, in the sense that more realistic filter shapes started to be assumed. The main improvement has been that a filter shape with finite passbands is able to explain psychophysical data in a better way (Schaeffer *et al.*, 1950; Patterson, 1974; Patterson, 1976).

Later on, characteristics of the auditory filter have been derived in several studies (see e.g. Patterson, 1976; Houtgast, 1977; Fidell *et al.*, 1983; Moore and Glasberg, 1983; Shailer *et al.*, 1990; Moore *et al.*, 1990; Rosen and Stock, 1992 ; Baker and Rosen, 2006). The methods used in these later studies are different or present improvements from the former methods used to estimate the CB, such as the band-widening and critical ratio approaches briefly described above. Thus, they have allowed more accurate descriptions of the auditory filter properties across frequency. A description of relevant methods that have been used to characterize frequency selectivity is given in section 2.

Several representations have been used to describe the characteristics of the shape and level dependence of the auditory filter. This has led to the existence different families of filters, all of which roughly agree in the general characteristics of the auditory filter, such as that it has a rounded tip, relatively sharp skirts, and shallow tails (see e.g. filter shape in figure 2 and description below). However, the representations vary in whether they describe only frequency domain characteristics, whether they provide as well a time domain representation, or if they assume or not a degree of level dependency.

The most common representation is the "rounded-exponential", so-called *roex* filter, which is a frequency-domain-based description of tuning (Patterson *et al.*, 1982). In this

type of filter, the attenuation of the skirts away from the tip is approximated by an exponential function. A popular version, the roex(*p*,*r*) filter, can be described by the following equation:

$$W(g) = (1-r)(1+pg)\exp(-pg)+ r \quad (2),$$

where $W(g)$ is the filter power weighting function, g is the deviation from the center frequency (CF) of the filter divided by the CF, p is a parameter which defines both the tip bandwidth and the sharpness of the slope of the filter, and r is a parameter which determines the dynamic range of the filter. The effect of the latter parameter is to flatten the skirts for large values of g. The basic form of this filter is illustrated in figure 2, together with a rectangular "critical band" filter for comparison.

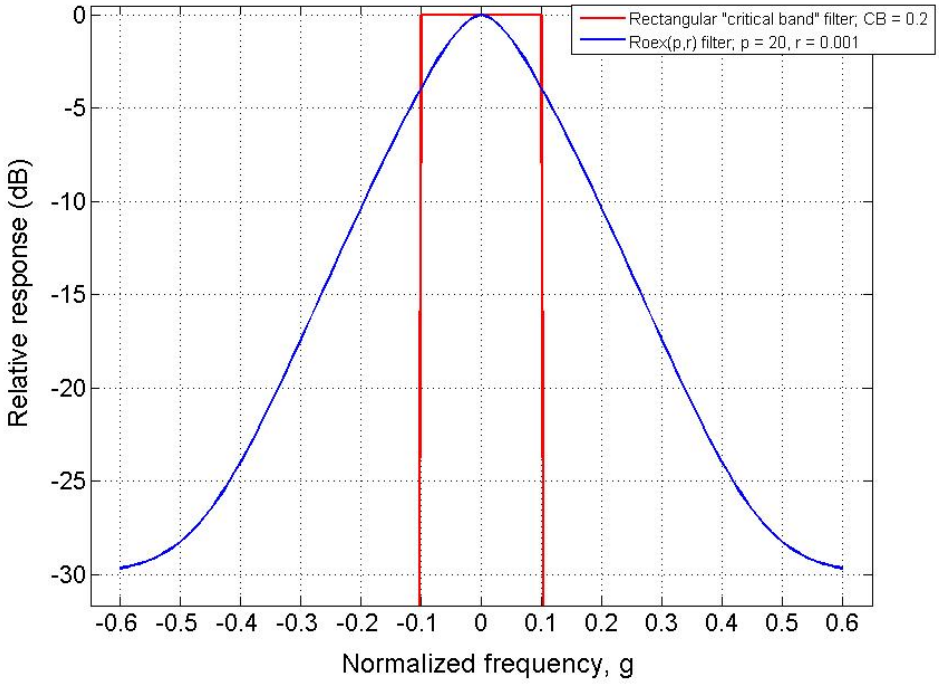

Figure 2: Power response of example symmetric roex(*p*,*r*) auditory filter (blue line, with parameters indicated in the legend). A rectangular "critical band" filter with CB equal to the effective bandwidth of the *roex* filter is shown for comparison (red line; see below in the text for details).

On the other hand, there is the family of the "gammatone" filters, which are a time-domain-based representation, in which the envelope of the impulse response is a gamma distribution and the fine structure is a tone. When it was first introduced it was used to describe cochlear impulse responses measured physiologically (Johannesma, 1972 ; de Boer, 1975). However, this function has been found to match well the frequency characteristics of auditory filter shapes derived psychoacoustically in humans, such as those of the *roex* filter function described above (Patterson *et al.*, 1987). For more details the reader is referred to de Boer (1975; 1978) and Aertsen and Johannesma (1980).

A further development has been the “gammachirp” family of filters, where the level dependence and its effect on the asymmetry of tuning have been approximated. This has been done by adding a chirp function into the carrier term of the gammatone filter (Irino and Patterson, 1997). These filters provide good correspondence to both psychophysical and physiological data. For more details the reader is refered to Irino and Patterson (1997; 2001) and Patterson *et al*. (2003). A comprehensive comparison of the gammachirp and the *roex* family of filters can be found in Unoki *et al*. (2006b).

A convenient description, analogous to the CB value in the earlier studies, has been to calculate the effective bandwidth of the auditory filter shapes. This has been called “equivalent rectangular bandwidth” (ERB) in the newer studies. The ERB of a given filter can be defined as equal to “the bandwidth of a perfect rectangular bandwidth filter which has a transmission in its passband equal to the maximum transmission of the specified filter and transmits the same power of white noise as the specified filter” (Moore, 1986). For the auditory filter shape shown in figure 2, ERB/f_0 ~ 4/p = 0.2, which is the same as for the “critical band” filter shown (i.e. CB/ f_0 = 0.2).

Based on data for CFs above 100 Hz, Glasberg and Moore (1990) proposed an equation relating the average ERB of the auditory filters for young listeners with normal hearing, denoted ERB_N, to CF:

$$ERB_N = 24.7 \times (4.37CF/1000+1) \quad (3)$$

Figure 3 shows the ERB values predicted by eq. (3) together with the CB values proposed by Scharf (1970) and results from example (more recent) studies that have included measurements at low frequencies. Above about 800 Hz, this formula gives ERB values roughly similar but somewhat smaller than CB values proposed by Scharf (1970). The latter values were largely based on results from Zwicker and co-workers (1957) as well as Greenwood (1961) and Hawkins and Stevens (1950) (the latter multiplied by 2.5 to account for their wrong estimate of *K*). However, at low frequencies the results especially differ, with the CB flattening off and reaching an asymptote value of 100 Hz below about 500 Hz. Contrary to this, the ERB function, in agreement with results from more recent studies, shows a continuous decrease in bandwidth at least down to a center frequency of 100 Hz, where it reaches a bandwidth value of about 35 Hz. Since the newer studies have perfomed more direct measures of frequency selectivity, unlike the older CB estimations, some of them based on the critical ratio, it is generally accepted that the bandwidth continues to decrease below 500 Hz at least down to CF = 100 Hz (Moore and Sek, 1995).

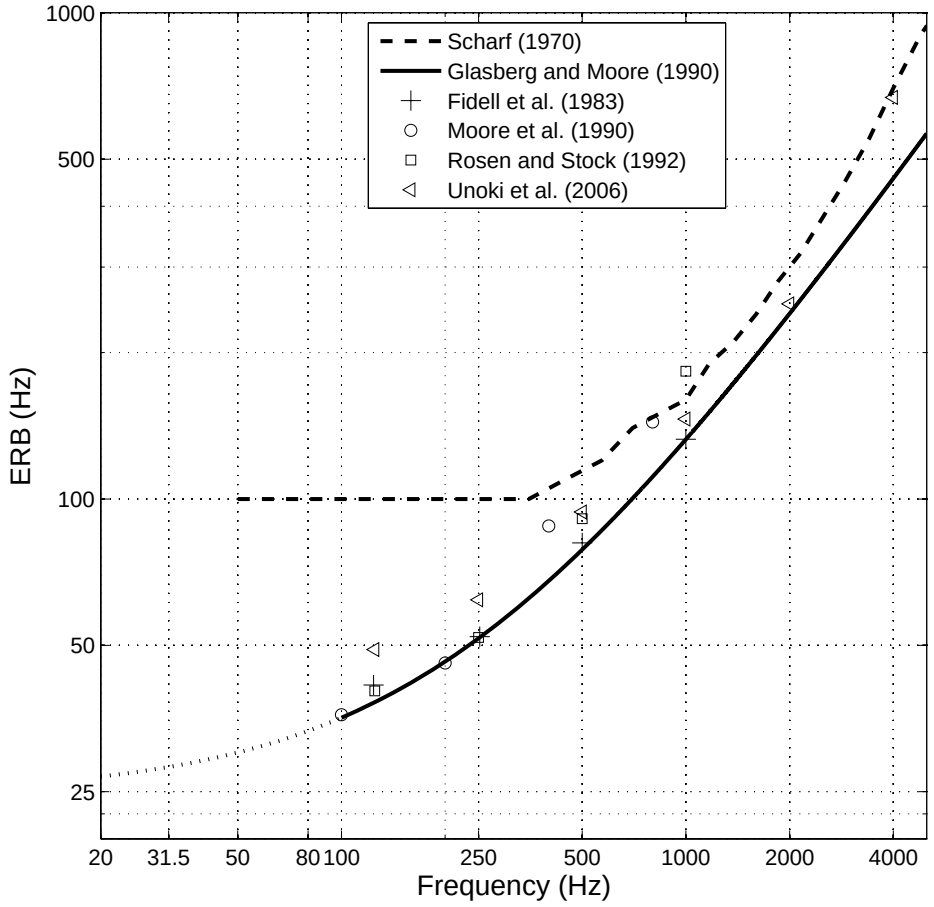

Figure 3: Equivalent rectangular bandwidth function (see equation 3) proposed by Glasberg and Moore (1990) and critical bandwidth values proposed by Scharf (1970) (solid line and dashed line, respectively). The dotted line shows extrapolated values below 100 Hz according to equation 3. The symbols show results from example studies that have included measurements at low CFs (indicated in the legend).

1.2 Further considerations regarding frequency selectivity and the perception of complex sounds

It is important to recall that listeners make in fact use of information across different auditory filters. The relative response of several auditory filters to a complex input sound has been used as an approximation of the internal representation that a sound creates in the cochlea, and appears to fairly account for what influences our perception (Patterson and Moore, 1986; Moore *et al.*, 1997). Even under a single sinusoidal input, the internal representation of such signal will be approximated by an *excitation pattern*, which is the output of different auditory filters plotted as a function of their CF (Moore and Glasberg, 1983). The amount of sound energy in different frequency bands is also the basis of models used to predict the detectability and loudness of complex sounds (Zwicker *et al.*, 1957; Zwicker and Scharf, 1965; Moore *et al.*, 1997). In very brief terms, what the

evidence from a variety of experiments has suggested is that the ear is capable of integration over bandwidhts much larger than the CB (see e.g. Spiegel, 1981; Buus *et al.*, 1986), and auditory models need to account for such effects.

In this respect, methods that aim to characterize frequency selectivity (described below), generally provide conditions where isolation of a single auditory filter is seeked, i.e. where the output from a single filter is probably dominant. Furthermore, some methods require the assumptions of the power spectrum model of masking, described in section 1.1. In this model, the masker is represended by its long-term spectrum, while phase effects and short-term fluctuations are ignored. These are only approximations, supported to some extent by defined experimental conditions and analysis tools. For example, when detecting a signal, listeners may in fact use a filter not centered at the signal frequency, but at another frequency that improves the signal-to-masker ratio (Patterson and Nimmo-Smith, 1980). This has been called “off-frequency listening”, and if it is not avoided or at least considered in the analysis, it is a violation of the power-spectrum model.

On the other hand, we are in fact sensitive to relative phases of components in complex sounds both within (Zwicker, 1952) and across CBs (Blauert and Laws, 1978; Patterson, 1987). Besides, regular amplitude fluctuation in the form of beats, as a result of interaction of signal and masker, can highly influence detection of a signal, especially when masker and signal are both sinusoids (Wegel and Lane, 1924; Egan and Hake, 1950; Greenwood, 1971; Alcántara *et al.*, 2000). This suggests that not only the amount of activity (Delgutte, 1996), but probably also the temporal pattern of neural firings elicitated by the signals (Carney *et al.*, 2002) play an important role in signal detection. While some of these factors, such as slow amplitude fluctuations in narrowband noise, may affect the efficiency in the detection process (Zwicker and Schutte, 1973; Moore and Glasberg, 1987) rather than directly the estimates of frequency selectivity, others, such as beats between signal and masker, may lead to wrong estimates of frequency selectivity if not controlled.

Another observation concerns the artificial segmentation of the auditory-filter-bank construct. Some results, especially from the older CB data, have been expressed as bandwidths at specified CFs, and without overlapping edges (see e.g. Scharf, 1970). This is a simple manipulation done for convenience, which may be misleading, since the hearing organ does not present such discontinuities. With the exception of the lower and upper limits where auditory filters may hypothetically be centered, imposed by the physiology of the hearing organ, the auditory-filter-bank construct should be generally regarded as a continuous set of overlapping bandpass filters.

2 Measuring frequency selectivity

Since the work of Fletcher (1940), a variety of psychoacoustic methods have been used to estimate the characteristics of frequency selectivity across frequency. Besides the band-widening and critical-ratio approaches briefly described in section 1.1, other relevant methods include: a) two-tone masking experiments; b) detection-of-complex experiments; c) rippled-noise method; d) notched-noise method; and e) psychophysical

tuning curves. In the following, a brief description and background of these methods is given. Methods d) and e), which were used in the experimental work reported in the main body of this thesis, are treated somewhat in more detail. Sub-section 2.6 describes general methodological challenges associated with the low-frequency range.

2.1 Two-tone masking

In this method two pure tones are used to mask a narrow band of noise. The noise band is placed midway between the two tones (of equal amplitude). This is in some manner opposite to Fletcher's approach described in section 1.1, where a noise is used to mask a tone. The two masking tones are then moved apart from each other in frequency, and threshold for the noise band is measured as a function of the frequency separation between the tones. This method was used by Zwicker (1954) to estimate values of the CB. He found that the masked threshold of the noise remained roughly constant up to a critical separation, after which threshold droped abruptly. The point at which the decrease began was taken as a measure of the CB.

Replications of this experiment in subsequent work have included variations where, instead of a noise band, a pure tone is used too as signal to be masked by the two surrounding tones. Results from these later studies, however, have not consistently shown abrupt threshold-drop transitions (Patterson and Henning, 1977; Glasberg *et al.*, 1984 ; Rabinowitz *et al.*, 1980). It has been therefore suggested that this method may be susceptible to audibility of distortion products (created by interaction between the lower tone and the noise band in the hearing organ) and that the abrupt transitions found by Zwicker may have been a consequence of that. If such effect is avoided (i.e. by masking the combination products) then the threshold-drop transition is much smoother and results are more in accordance with realistic filter shapes with finite passbands, derived for example from using notched-noise or rippled-noise maskers (see below). Unlike for the band-widening approach or the use of notched-noise maskers, an energy detection model such as that described in eq. (1) was not able to predict absolute masking levels obtained with this method (Patterson and Henning, 1977).

2.2 Threshold of complex sounds

In this type of experiment threshold of a complex sound, composed of multiple sinusoids (i.e. partials), is measured as a function of the number of partials and the frequency separation between the two partials at the extremes, i.e. those with lowest and highest frequencies. These experiments make use of the observation that when several sinusoids are presented together they may be audible, even when each component in isolation is below absolute threshold. An alternative version of this experiment considers measuring the detection of a noise band as a function of its bandwidth.

Gässler (1954) estimated the CB using this method. His procedure was as following. At first, as a preliminary setting, uniform masking noise was used to make the audiogram flat over a wide frequency range. Then, threshold for a single tone was measured. Afterwards, another tone spaced 10 Hz below was added and threshold for the 2 tones

was determined. As expected if their powers were added, the levels of each of the 2 tones at threshold was lower than for 1 tone in isolation. He continued this procedure of adding tones 10 Hz apart (this was done for various tone frequencies), and found that the levels of each component required to achieve threshold continued to decrease, making the total energy required for threshold roughly constant, up to a critical point, after which the total energy required for threshold increased. He took this point as a measure of the CB, suggesting that when components are distributed whithin a CB their energies are summed, and threshold is determined by their summed energy, but that detection was less good when components were distributed over more than a CB.

Resuts from more recent studies, however, do not agree completely with these findings. These studies have shown that the hearing organ is capable of integration over bandwidths much larger than the CB (Spiegel, 1981; Buus *et al.*, 1986; Langhans and Kohlrausch, 1992). Therefore, this approach seems to provide information of the functioning of not only one, but of several auditory filters, and seems unsuitable if isolation of a single filtering unit is seeked.

2.3 The rippled noise method

Houtgast (1977) used a rippled noise masker to obtain auditory filter shapes for simultaneous and non-simultaneous masking. Rippled noise has a spectrum level that varies sinusoidally on a linear frequency scale, and is constructed by delaying a random noise by an amount (typically a few miliseconds) and adding it back to the original. In his procedure, he fixed the level of a sinusiodal signal and obtained masked threshold for the rippled-noise by varying its mean spectral level, i.e. the spectral level at half way points between a peak and a trough in the noise. Different parameters of the noise, such as peak-to-peak distance, were varied and threshold was determined in each condition. Given that the tone level was fixed, the threshold data describes how the mean spectrum level of the noise must be varied to achieve a constant noise level at the output of the auditory filter. By modeling the auditory filter as a Fourier series and assuming a power spectrum model of masking (see eq. (1), section 1.1), Houtgast was able to extract the Fourier coefficients of the auditory filter from the corresponding rippled-noise threshold levels. From this procedure he found auditory filters with rounded tips and sharp skirts, broadly similar to those found using notched-noise maskers (Glasberg *et al.*, 1984; Niemiec and Yost, 1992). This method does not allow, however, to measure the filter shape over a wide dynamic range. Besides, rippled noise gives a particular pitch sensation, which makes it more difficult for the subjects, as they cannot only be instructed to listen to a “tone in the noise”, such as with the notched-noise method. More details of this procedure can be found in Patterson and Moore (1986).

2.4 The notched-noise method

In this method, a pure tone is positioned inside the spectral notch of a noise masker (the latter is termed a “notched noise”) and masked threshold for the tone is measured considering different notch widths (de Boer and Bouwmeester, 1974; Patterson, 1974; 1976; Patterson and Nimmo-Smith, 1980). A variation of this method considers fixing the

level of the tone and varying the level of the noise to obtain masked threshold for the latter (Rosen *et al.*, 1998; Baker and Rosen, 2006). Figure 4 shows an illustration of the method. By placing the tone in a spectral notch, off-frequency listening is controlled. The method allows also to measure the asymmetry of tuning, by placing the flanking bands asymmetrically around the tone.

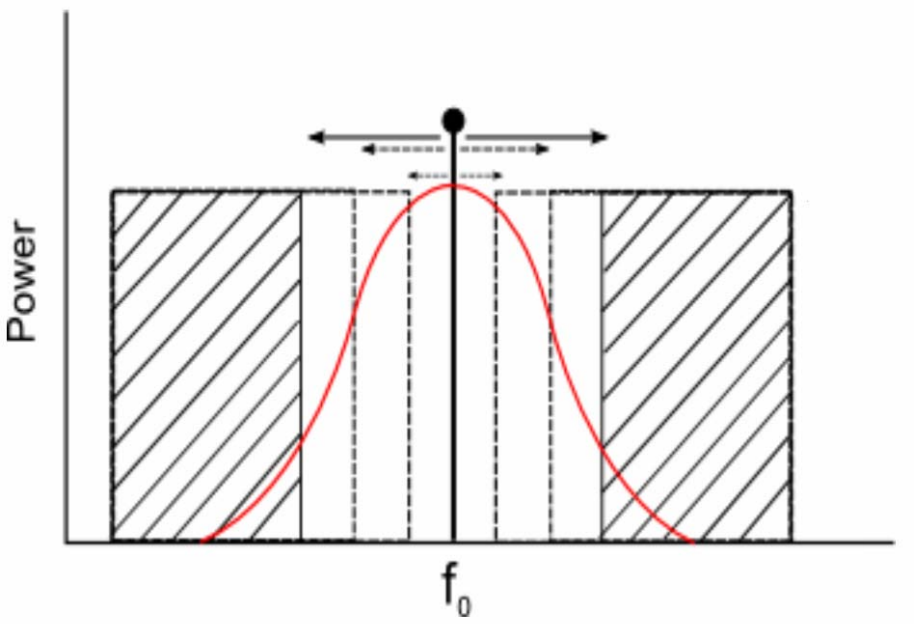

Figure 4: Spectral configuration of the signal (tone with frequency f_0, shown in vertical-solid line) and maskers (flanking bands shown in shaded-area boxes) in a notched-noise experiment. The arrows indicate how different notch-widths are arranged to cover different regions of the auditory filter (the latter is shown in red line).

Assuming a power spectrum model of masking and using very sharp spectral edges for the flanking bands, eq. (1) becomes:

$$P_S(g) = Kf_0N_0\left(\int_{-\infty}^{-g} W(g)dg + \int_{g}^{\infty} W(g)dg\right) \qquad (4),$$

where $W(g)$ is the auditory filter power weighting function and g is the normalized deviation from the CF.

Since threshold is measured for several notch widths (i.e. values of g), each threshold representing the amount of noise that leaks under the filter function (for a noise with constant spectral density, N_0, the latter can be taken out of the integrals), the auditory filter shape can be obtained by taking the derivative of the threshold curve as:

$$W(g) = -\left(\frac{1}{2KN_0 f_0} \frac{dP_S}{dg} \right) \quad (5),$$

where symmetry has been assumed for simplicity. In this form of the procedure no specific shape for the filter needs to be assumed. However, if specific filter shapes are assumed, such as filters from the *roex*, gammatone, or gammachirp families, then the integral can be directly evaluated and the parameters of the filter fitted to the masked-threshold data. An example of how eq.(4) would look for a roex(*p,r*) filter (see eq. (2), section 1.1) is shown in eq. (6).

$$P_S(g) = 2KN_0 \left[-(1+r)p^{-1}(2+pg)e^{-pg} + rg \right] \Big|_g^L \quad (6),$$

where L is an integration limit representing how far away from the CF any contribution of the lower/upper edge of the lower/upper flanking band is considered, typically set to 0.8 as recommended by Patterson (1982).

In a similar way as for the rippled-noise method, assuming the general masking model of eq. (1), the efficiency in the detection process can be separated from frequency selectivity. This allows these methods to more directly estimate the characteristics of frequency selectivity, unlike approaches such as the critical ratio (see section 1.1). The notched-noise method is also more sensitive in describing the overall filter shape than the band-widening approach. This is because for the latter, the contribution (in masking the signal) of spectral components of the noise in the region where g is around a CB is negligible. Therefore, the filter tails cannot be described. The notched-noise method, instead, allows a better definition of the filter skirts and tails using the large notch-width conditions.

2.5 Psychophysical tuning curves

A psychophysical tuning curve (PTC) is the level of a masker, L_m, required to just mask a signal which is fixed in frequency and level, plotted as a function of the frequency of the masker, f_m. It is the psychophysical analogous of neural tuning curves measured in the auditory nerve (see e.g. Zwicker, 1974; Pickles, 1984; Narayan *et al.*, 1998). The level of the signal is kept relatively low, in order to excite a small number of neurons with characteristic frequencies close to that of the signal. Figure 5 shows an illustration of a PTC.

PCTs have generally been found to have similar features as neural tuning curves measured physiologically (Small, 1959; Vogten, 1974; 1978; Moore, 1978). Since the level of the input (i.e. masker level) is varied while the output level (i.e. signal level) is fixed, they approximate an inverted auditory filter shape, assuming linearity. When sinusoids are used as both signal and masker, beatings can occur (Wegel and Lane, 1924; Egan and Hake, 1950; Greenwood, 1971; Alcántara *et al.*, 2000), an effect not associated with frequency selectivity which greatly affects detection of the signal. The use of beats

as a cue can be reduced by using narrowband noise as a masker (Moore *et al.*, 1998; Kluk and Moore, 2004), or using "beating tones", that beat at the same rate as the signal and masker, via the phenomenom of modulation-detection interference (Yost *et al.*, 1989; Bacon and Moore, 1993; Alcántara *et al.*, 2000).

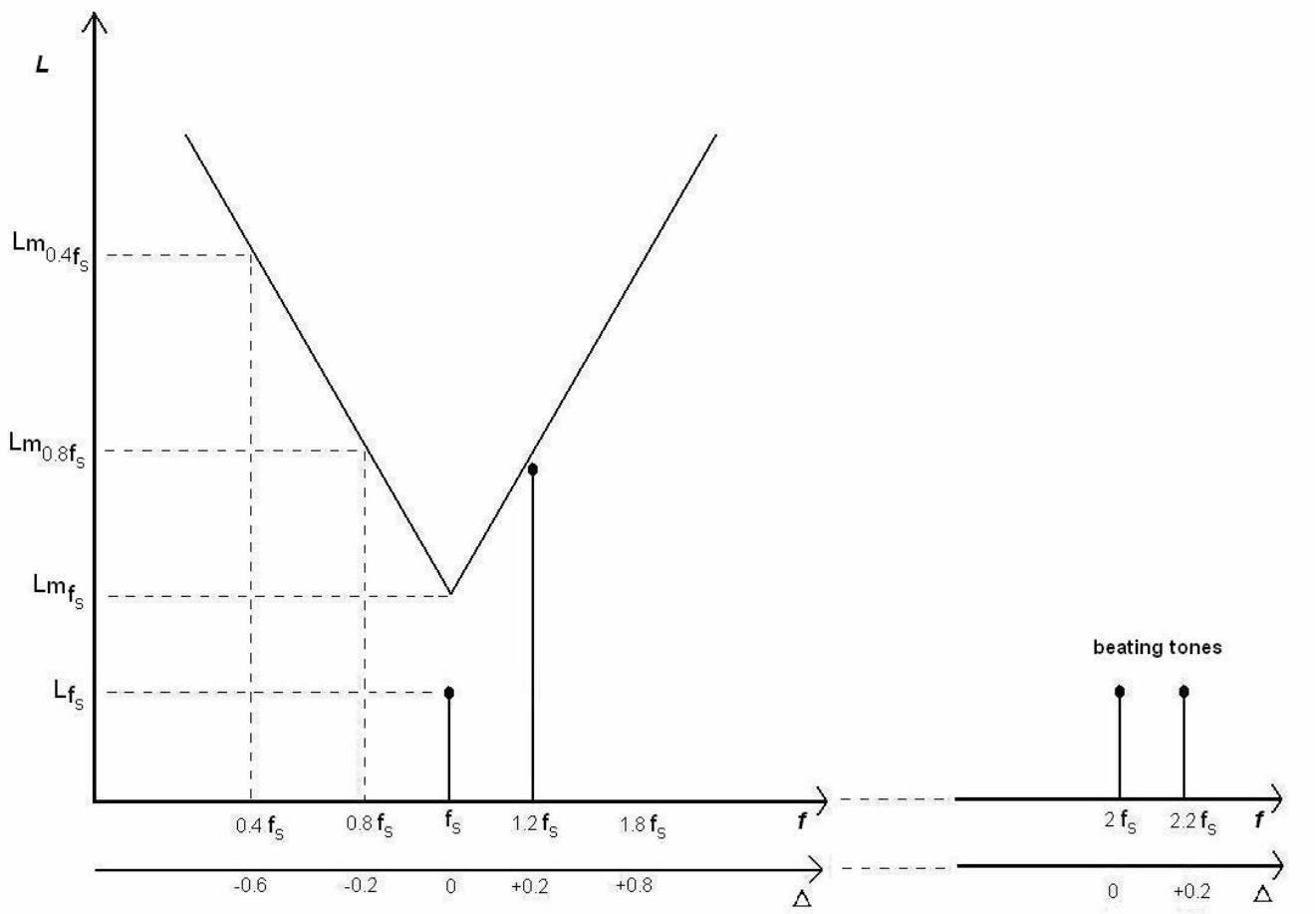

Figure 5: Schematic representation of signal and masker in a PTC experiment. The signal has frequency f_s and the masker assumes different frequencies; in each case the level (L_m) to just mask the signal is found, as shown in vertical-dashed lines and vertical-solid line. The lower axis shows the frequency separation between signal and masker in normalized terms, with $\Delta = \left(f_m - f_s / f_s \right)$. On the right side, an example of how beating tones may be configured to beat at the same rate as the signal and masker is given (signal, masker, and beating tones in the example are shown in vertical-solid lines).

An example of how beating tones may be configured to avoid using beats as a cue is also given in figure 5 (right side). The beating tones are positioned around an octave above the signal to avoid direct masking effects. On the other hand, off-frequency listening can

occur (Johnson-Davies and Patterson, 1979; O'Loughlin and Moore, 1981), especially when the signal and masker are close in frequency. To avoid this effect, a low-level background noise with a spectral notch is sometimes used (O'Loughlin and Moore, 1981; Moore *et al.*, 1984). Within its constrains of assuming linearity, a relative advantage of this method is that it offers a direct approximation of the auditory filter shape, without assuming a specific model to derive it.

2.6 General challenges associated with the low-frequency range

An important factor to take into account when performing psychoacoustic experiments to characterize aspects of low-frequency hearing is the large variation of hearing sensitivity with frequency in the low-frequency range. Sensitivity to sound drops abruptly at low frequencies, particularly below 100 Hz, as the large steepness in absolute thresholds in this range reflects (see e.g. ISO 226, 2003). Since the above-described methods require a significant excitation of the hearing organ to achieve, for example, sufficient masking of a signal by a masker, this sets particular demands on different aspects of the problem. Principal aspects are: the setting of proper experimental conditions –which will depend on the method used– ; the subsequent data analysis and interpretation; and the demands on the sound reproduction system used. These are described in the following.

2.6.1 The setting of adequate experimental conditions

The design of the experiments should be such that over all –or at least the most relevant– configurations of signals and maskers used, it is ensured that the masker produces sufficient masking over the signal. This should be so even when the masker has frequency components well below the signal frequency, condition in which the signal would normally be much more audible than the masker, due to e.g. the high steepness in low-frequency absolute thresholds. In an approach such as the notched-noise method, used in part of the experimental work described in this thesis, this was achieved by providing a (sufficient) gain to the lower flanking band relative to the upper band (i.e. providing emphasis to the lower flanking band). When psychophysical tuning curves were measured, the masker was adaptively varied in level, providing the necessary (relatively high) levels required to mask the signal in the most extreme conditions.

2.6.2 Considerations regarding the analysis and interpretation of results

The large decrease in hearing sensitivity at low frequencies also demands specific attention from the point of view of the analysis and interpretation of data. Particularly, the factors behind this sensitivity drop must be carefully treated. For example, unlike at higher frequencies, where the audiogram is much flatter, characteristics of the at-ear spectrum of the stimuli probably do not realistically reflect the characteristics of the spectrum that reaches the cochlea. The latter can be better approximated by passing the stimuli through a filter that represents the frequency dependent filtering produced by the middle-ear, which has a significant effect at low frequencies (Puria *et al.*, 1997; Aibara *et al.*, 2001). In this context, it is important to note that while the displacement of the middle-ear ossicle chain is stiffness dominated below about 500 Hz –providing an about

+6 dB/oct slope–, the middle-ear-transfer function increases largely in slope below about 50 Hz (Marquardt *et al.*, 2007; Marquardt and Pedersen, 2010). This is because this transfer function includes the effects of the helicotrema shunt mechanism, which equalizes the pressure across the cochlear ducts (see e.g. Marquardt and Hensel, 2008). Therefore, even if part of the masker is emphasized, due to the strong influence of the middle-ear-transfer function, a proper analysis should account for its filtering effects. Besides, the fact that the helicotrema shunt is a cochlear mechanism demands special treatment of at least part of the middle-ear-transfer function. The latter can conceptually no longer be treated as being a fixed filter applied prior to (cochlear) auditory filtering, as usually done at higher frequencies. Therefore, the region below about 40 to 50 Hz in the middle-ear-transfer function, where the helicotrema shunt appears to be dominant, has been treated with special attention in the analysis of all experimental data. As well, to perform more specific analyses, individual estimates of this transfer function were obtained from both equal-loudness contours and a distortion-product-isomodulation technique.

Another aspect important to consider is the dependence of signal detection on the time pattern of the stimuli. It is well known that in a large portion of the audible frequency range –below about 5 KHz–, information about input sound is partly conveyed in the temporal pattern of nerve spikes, which can phase lock in response to a stimulus waveform. Since at low frequencies, considering a fixed signal duration, relatively less stimulus cycles will occur than at higher frequencies, factors such as the efficiency in the detection process may be affected. It is therefore convenient to use generally longer signal duartions (i.e. above 0.5 to 1 sec.) at low frequencies than durations usually used at higher frequencies. The stimuli should also have sufficiently long ramps to avoid spectral splatter effects at the onset of the signals. Further, since low-frequency maskers usually will have a narrowband spectrum, the stimuli will present inherent slow amplitude fluctuations which may also have effects on signal detection. From this perspective, a useful condition, if random noise is used as stimulus, is that it contains representative amplitude fluctuations. This can be done by either using a continuous noise or by extracting samples from a (representative) long duration buffer (i.e. >30 sec.); the latter can be done for practical purposes, if signal detection under both conditions is similar. On the other hand, the effect of amplitude fluctuation in the form of beats, due to interaction between low-frequency signals and maskers, is also important to evaluate in order to determine how much it may affect the estimates of tuning.

In general, the use of an approach such as the notched-noise method is thought to be advantageous, since the efficiency in the detection process can be separated from auditory tuning. However, since the derivation of auditory filter shapes from this approach usually relies on the assumptions of the power-spectrum model of masking, it was considered necessary –for comparison, validation of main findings, and extension of the results– to include another methodology in the experimental work, such as the measurement of psychophysical tuning curves. The latter can give more direct (i.e. model independent) estimates of frequency selectivity, and are thought to be useful for measurements that comprise very low frequencies, since there may be regions where no tuning occurs and these will be more directly identified.

2.6.3 Considerations regarding the reproduction of low-frequency sound signals and the equalization of the sound system

Finally, an always present demand when controlled playback of low-frequency sound stimuli is needed relies upon the sound reproduction system. When playing back low-frequency acoustic signals, a common limitation are distortions produced by the loudspeaker(s). These arise due to inherent physical limitations in the transducers, and can be classified into linear and non-linear distortion. The former refer to particularities in the frequency response of the loudspeaker(s), which at sufficiently low frequencies will take the form of a rolling off in the response. This happens due to two principal factors. On the one hand, for a given diaphragm radius, at low frequencies the wavelength becomes comparable or larger than this radius, and the real part of the radiation impedance (associated with power radiation) will be comparatively smaller than its imaginary part. A a result, radiation efficiency will be poor at low frequencies. On the other hand, the mechanical impedance of a loudspeaker will increase at low frequencies, due to the stiffness of the suspension system; the latter will also set a limit for the mechanical resonance frequency of the transducer.

On the other hand, non-linear distortions, such as harmonic or intermodulation distortion, produce spectral components which are not present in the driving input signals. Among the general sources of non-linearities in electrodynamic loudspeakers are non-linearities in the stiffness of the suspension system and in the force factor acting over the voice coil; both non-linearities depend highly on the displacement of the voice coil; with the force factor non-linearity depending as well on the current. Because of the generally inefficient sound radiation at low frequencies, relatively larger input signal levels will be typically required, and as a result nonlinearities are likely to occur. For example, when system non-linearities add harmonics of e.g. a pure-tone (fundamental) input signal, the relative level of the harmonics will generally tend to increase when the fundamental frequency is low, due to the relatively more efficient radiation of the harmonic frequencies by the loudspeaker(s). For more details about loudspeaker non-linearities the reader is referred to Klippel (2006).

In listening tests, both linear and non-linear distortions produced in the playback of sound stimuli can cause problems if not controlled. If the linear response of the loudspeaker(s) is not accounted for, this may lead to wrong conclusions about the relation between perception and the physical characteristics of the stimuli. Non-linear distortions can produce the same latter effect if not controlled, due to audibility or interference effects produced by the creation of additional spectral components.

These problems can be partially overcome by a combination of factors. For example, the use of loudspeakers with relatively large diaphragms will improve low-frequency sound radiation efficiency. As well, if the suspension system is made very compliant, the mechanical impedance and resonance frequency can be lowered. However, there must be a trade-off since if the suspension stiffness is too low, the voice coil may move into non-uniform regions of the magnetic field, and increase non-linear distortions.

For a given setup, conditions can be optimized by combining a dedicated equalization of the loudspeaker(s) and, within the constrains imposed by system non-linearities and the requirements of the psychoacoustic method, limiting the overall levels of the stimuli to be used. An adequate equalization of a sound-reproduction system can be achieved by designing digital filters. The latter should be based on measurements of the frequency response of the system, done preferably in high resolution (i.e. obtained for several closely-spaced frequencies). These measurements should be performed in low background-noise conditions and provide enough signal-to-noise ratio for the lowest frequencies aimed for equalization.

In the main experimental work two different setups were used. In one case, due to the relatively large frequency range necessary to cover, headphones were used as transducers. In this case, the responses were individually measured (with a microphone positioned flush at the ear-canal entrance) to account for observed individual differences in the at-ear responses. The average response of several re-positionings was obtained to

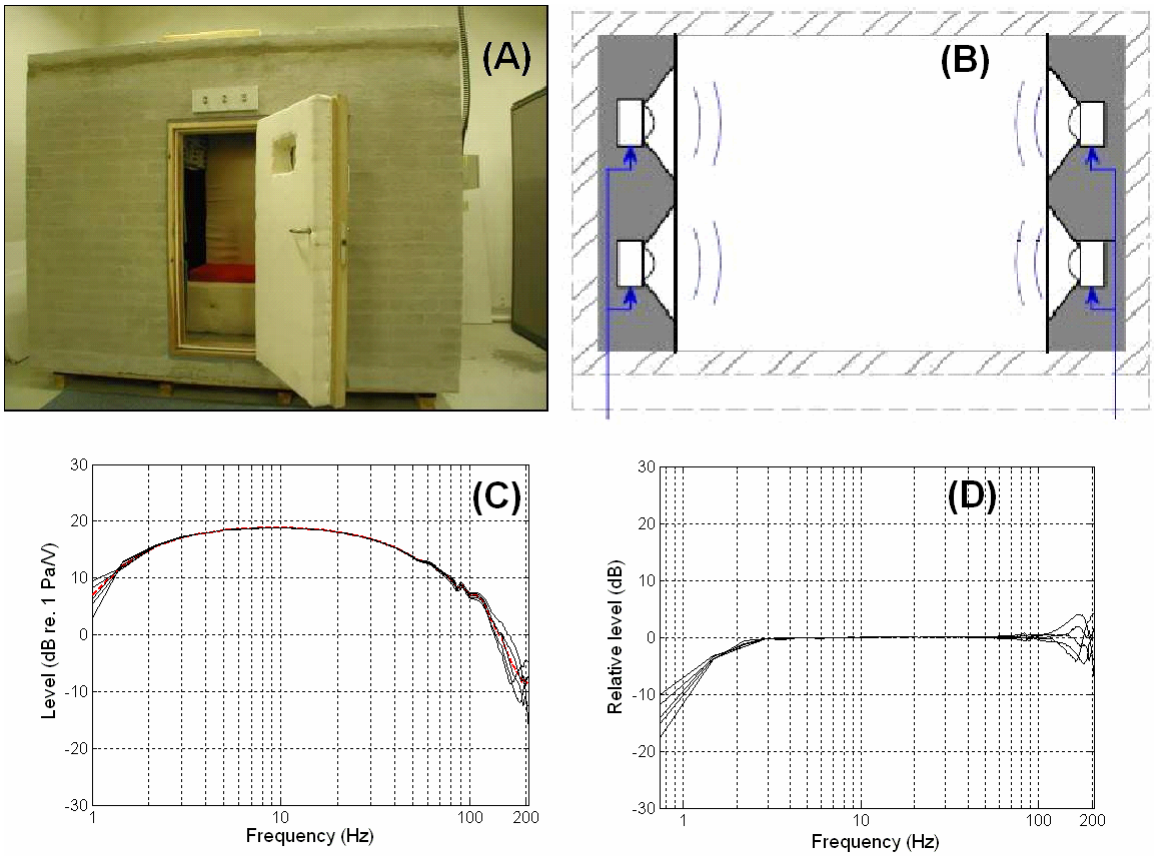

Figure 6: (A) Pressure-field chamber exterior view; (B) Front view diagram where each inner volume can be observed. The 2 loudspeaker columns are doubled (i.e. there are 4 loudspeakers in each side wall); (C) Frequency response of the pressure field chamber (including amplification system) at 5 listening points (solid-blak lines) and average response (red-dashed line); (D) Equalized responses after passing each of the 5 listening-point responses through a digital filter designed to match the inverse of the average frequency response.

possibly account for changes with positioning. The individually designed digital filters approximated the inverse of the measured frequency response. However, excessive gain was not allowed, particularly at the lowest frequencies where the response of the transducer naturally rolled off, in order to avoid an increase in non-linear distortions. The amount of of non-linear distortion, particularly harmonic distortion was thoroughly evaluated. From these measurements, adequate target levels for the stimuli were chosen so as to minimize the amount of non-linear distortion while, at the same time, provide control of the acoustic signals by means of the digital-filter equalization.

A similar overall approach –although individual calibration was not required– was used in the other main setup used in the experimental work, which consisted of a "pressure-field chamber". This setup had been previously constructed especially for psychoacoustic experiments involving low-frequency sound. It was designed to maximize low-frequency performance, allowing controlled playback of high sound-pressure levels at low frequencies, while providing exposure conditions where the sound pressure is uniform within a volume –inside which the listener is located. A picture of the chamber is shown in figure 6 (panel A). It is constructed of concrete floor and ceiling, and bricks walls covered with cement. It basically consists of 3 volumes: two side volumes (rougly 0.7 m^3 each) which act as loudspeaker enclosures, and a center volume of about 1 m^3 (see panel B). The latter volume is occupied by human subjects and was designed as a trade-off between being not overly uncomfortable and providing a pressure field up to the highest possible frequency.

The chamber is equipped with four Seas 33 F-WKA 13-inch (bass) loudspeakers in each side wall, positioned behind covering panels. The listeners sat facing the door. The frequency response of the room is shown in panel C, considering 5 listening positions where it was expected that the listeners would be located. Even though the response shown in panel C is relatively smooth, it was compensated by a digital filer designed to match the inverse of the average frequency response. On panel D, each of the 5 measured responses are shown after equalization. Although the cabin provides an effective pressure field within its overall volume up to about 61 Hz, level differences in space are below ±3 dB up to about 150 Hz within an expected listening space. Harmonic distortion and background noise levels were measured and were found satisfactory for the purpose of the experiments. More details of the setup are given in the main body of the thesis (see Manuscript 2).

3 Organization of the thesis

The main body of the thesis are four manuscripts that describe a series of experiments carried out during the PhD period. Each is an independent work. However, they are related to each other, as described in sub-section 3.1. The manuscripts are the following:

Manuscript (1): Jurado, C. A. and Moore, B. C. J. (2010). "Frequency selectivity for frequencies below 100 Hz: comparisons with mid-frequencies," *Journal of the Acoustical Society of America*, Vol. 128, issue 6, (in press).

Manuscript (2): Jurado, C. A., Moore, B. C. J. and Pedersen, C. S. (2010). "Psychophysical tuning curves for frequencies below 100 Hz," submitted for possible publication in the *Journal of the Acoustical Society of America*.

Manuscript (3): Jurado, C. A., Pedersen, C. S. and Marquardt, T.M. (2010). "Frequency selectivity at very low centre frequencies: the influence of the helicotrema on individual differences in low-frequency sound perception," in *Proceedings of the 14th International Conference on Low Frequency Noise and Vibration and its Control*, Aalborg, Denmark, pp. 129-145.

Manuscript (4): Jurado, C. A. and Marquardt, T. M. (2010). "The effect of the helicotrema on low-frequency hearing: II. Equal-loudness contours," in preparation for submission.

The last manuscript has a companion manuscript which has not been included in this thesis. It is titled: "The effect of the helicotrema on low-frequency hearing: I. Human forward-middle-ear transfer functions", with Torsten Marquardt as 1st author and the author of this thesis as 2nd author.

3.1 Description of the manuscripts and interrelations

Manuscript (1): "Frequency selectivity for frequencies below 100 Hz: comparisons with mid-frequencies"

In this study, characteristics of frequency selectivity were obtained for center frequencies between 50 and 1000 Hz using the notched-noise method. This was done to compare main properties of the auditory filter at low frequencies with characteristics found at higher frequencies, and at the same time extend the frequency range where results existed.

The main findings were that, although the relative sharpness of tuning worsens with decreasing frequency, the bandwidth of the auditory filter decreases at least down to a center frequency of 80 Hz. Below that, the bandwidth increased with decreasing center frequency, the contrary of what is typically observed in most of the frequency range, reflecting that the mechanism of tuning was in some manner affected. Reasons for this worsening of tuning are investigated in detail in Manuscript (3). Figure 7 re-plots the

ERB values previously shown in figure 3 (see section 1.1) together with the ERB values obtained in this study –considering filter shapes that include the combined effects of all filtering processes; see solid-blue line and asterisks.

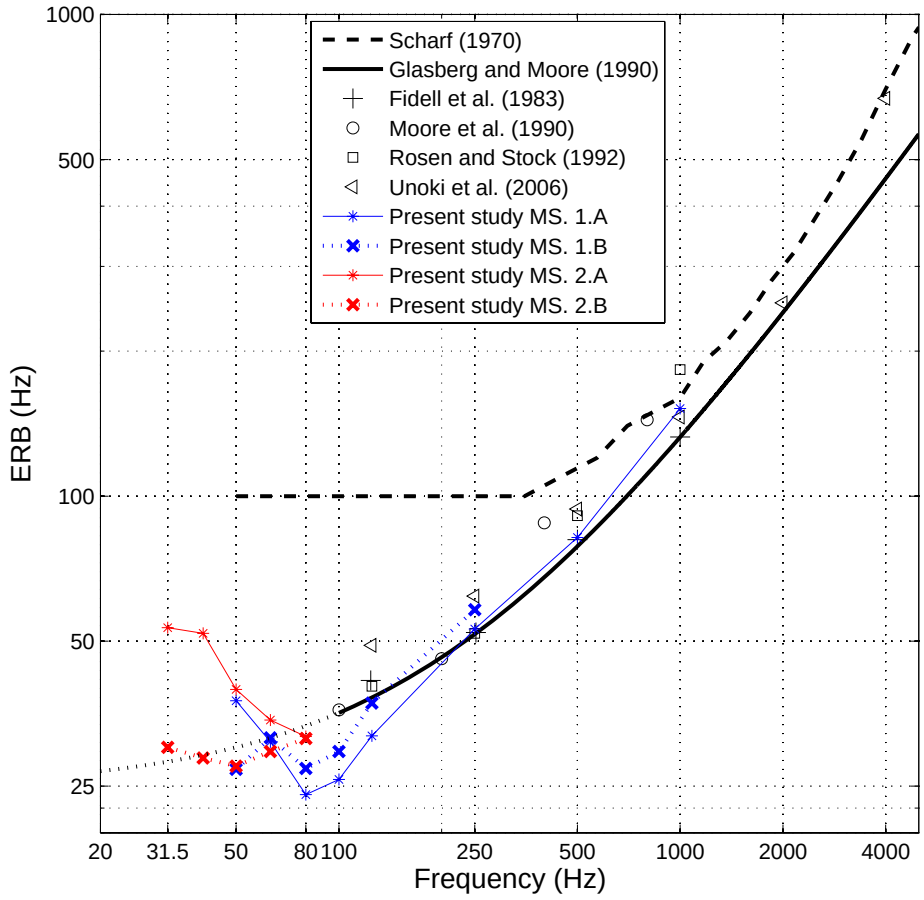

Figure 7: As figure 3 (see section 1.1), but mean ERB values from the present work have been included. Results described in Manuscript 1 (MS. 1) are shown in blue lines and symbols; 1.A: values obtained considering the combined filtering effects of the (assumed) middle-ear-transfer function and the derived filter shapes; 1.B: values obtained after including the middle-ear-transfer function as part of frequency selectivity below 50 Hz, but excluding it from frequency selectivity elsewhere. Results described in Manuscript 2 (MS. 2) are shown in red lines and symbols; 2.A: ERB values calculated for the mean psychophysical tuning curves; 2.B: ERBs of mean tuning curves; these were obtained after including the estimates of the middle-ear-transfer function as part of frequency selectivity only within the region of dominance of the helicotrema shunt.

Regarding the asymmetry of tuning, steeper lower skirts than upper skirts were observed below 100 Hz, also the opposite of what is typically found at higher frequencies. The relative effect on tuning of the assumed shape of the middle-ear-transfer function, which includes the effects of the helicotrema shunt mechanism, was largest at the lowest center frequencies considered, especially influencing the sharpness of the low-frequency skirt of the auditory filter. Assumptions regarding whether the middle-ear-transfer function (or part of it) should be included or excluded from frequency selectivity had noticeable

effects also on the trends observed in the ERB values. This is exemplified in figure 7, where the ERB values roughly flatten off below 100 Hz –consistent with the extrapolation– if the middle-ear-transfer function is included as part of frequency selectivity below 50 Hz, but excluded from it otherwise (see dotted-blue line and crosses). It should be noted that this is a manipulation that excludes the effect of the middle-ear-transfer function on tuning in a region where it otherwise appears to be still highly influential –a region between about 40 to 100 Hz; this is described in Manuscript (3).

On the other hand, the efficiency in the detection process was found to moderately improve in the frequency range below 100 Hz. This may have been a consequence of the ability of subjects to "listen in the dips", taking advantage of the slow amplitude fluctuations present in the narrowband noise maskers used.

Manuscript (2): "Psychophysical tuning curves for frequencies below 100 Hz"

In this work, a series of experiments aimed to characterize auditory tuning for frequencies below 100 Hz was performed. In the main experiments, psychophysical tuning curves were measured for sinusoidal signals between 31.5 and 80 Hz, considering different masker types.

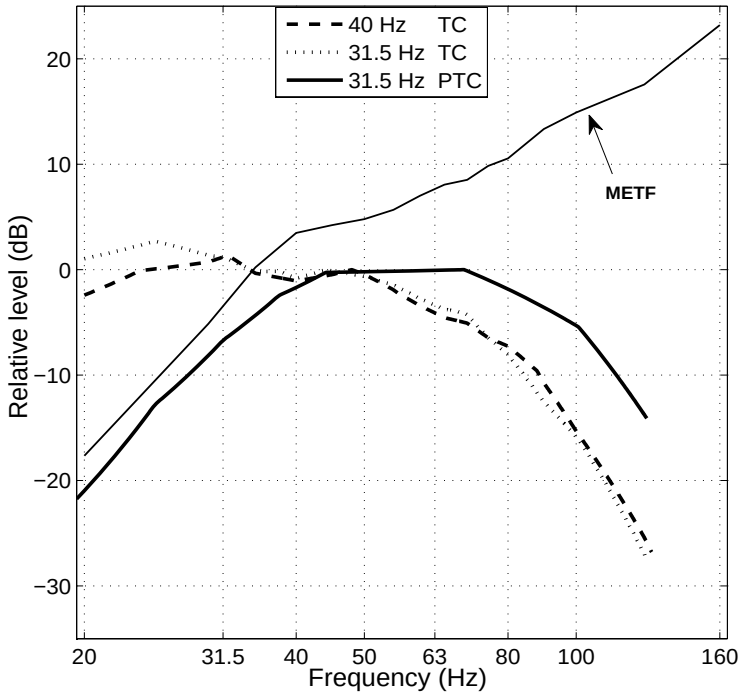

Figure 8: Mean psychophysical tuning curve obtained for a 31.5 Hz sinusoid (PTC; thick solid line) and mean tuning curves obtained for 31.5 and 40 Hz (TCs; dotted and dashed lines, respectively). The curves have been inverted and normalized by their maximum gain to resemble filter shapes. The tuning curves were obtained after completely excluding the (estimated) filtering effects of the middle-ear-transfer function (METF) from the shape of the PTCs. All curves are based on the results obtained using noise maskers. Further descriptions are given in Manuscript (2).

The focus was to obtain a detailed description of frequency selectivity at very low frequencies, extending the work described in Manuscript (1), and using a setup and method that increased the flexibility in experimental conditions and facilitated the interpretation of results. An equal-loudness contour was measured as well for each subject, to obtain a rough estimate of each subject's middle-ear-transfer function.

Main results agreed with the findings described in Manuscript (1), although now it was possible to extend the observations down to lower frequencies. In terms of the ERB, a similar trend of a worsening of tuning below 80 Hz was observed. This is shown in figure 7; see solid-red line and asterisks. Similarly as well, if the middle-ear-transfer function was included as part of frequency selectivity in the region of dominance of the helicotrema shunt, but excluded from it elsewhere, the values flatten off (see dotted-red line and crosses), consistent with the extrapolation shown in figure 7.

On the other hand, specific effects, such as upward-shifts in the tips of the tuning curves and large flattening of the upper skirts, could be distinctly observed. This is exemplified in figure 8, where the (inverted) psychophysical tuning curve for 31.5 Hz has been plotted. Clearly, its tip is well above 31.5 Hz and it presents a very flat upper skirt. These observed effects were in accordance with predictions of the possible influence of the middle-ear-transfer function on tuning, described in Manuscript (3). On the other hand, when sinusoids were used as masker and signal, the influence of beats was found to be roughly similar to what has been found at higher frequencies, with a marked effect on signal detection when masker and signal were relatively close in frequency.

The fact that for the lowest signal frequencies tested the tuning curves were very similar, and the use of equal-loudness contours to roughly estimate the shape of the individual middle-ear-transfer functions, allowed to resolve the center frequency of the "bottom" auditory filter (i.e. the one with the lowest center frequency). This auditory filter appears to be located between 40 and 50 Hz. Signals below this frequency would be detected via the low-frequency skirt of this filter. This is illustrated in figure 8, where together with the psychophysical tuning curve measured for 31.5 Hz, the tuning curves (i.e. the modified shape of psychophysical tuning curves after excluding the estimated filtering effects of the middle-ear-transfer function from tuning) obtained for 31.5 and 40 Hz have been plotted.

Manuscript (3): "Frequency selectivity at very low centre frequencies: the influence of the helicotrema on individual differences in low-frequency sound perception"

In this paper, the influence of the shape of the middle-ear-transfer function on frequency selectivity was examined. Individual differences in the shape of auditory filters derived psychophysically were compared with those obtained from predicting the possible effect of individual middle-ear-transfer functions on tuning (the latter transfer functions were based on objective estimates). The analysis described in this manuscript allowed to explain observations detailed in Manuscripts (1) and (2). In brief, results suggested that tuning is affected for auditory filters that would fall in a frequency region between about 40 to 100 Hz, due to distinct irregularities in the shape of the middle-ear-transfer functions in this region. These irregularities have the effect of flattening the upper skirt,

explaining the increases in ERB described in Manuscripts (1) and (2). The modeling allowed as well to predict upward shifts in the tips of tuning curves, which were observed in the experiments described in Manuscript (2). Due to individually-different particularities in the shapes of the middle-ear-transfer functions, individual differences in tuning were expected to generally increase with decreasing frequency, as the psychophysical results described in Manuscript (2) indicated.

Manuscript (4): "The effect of the helicotrema on low-frequency hearing: II. Equal loudness contours"

In this manuscript a comparison between the shape of equal-loudness contours and that of objective estimates of the middle-ear-transfer function is given. For a group of subjects, both equal-loudness contours and distortion-product-isomodulation curves were obtained and compared. The latter were assumed to reflect the shape of the middle-ear-transfer function, including the filtering effects of the helicotrema shunt mechanism (a cochlear mechanism).

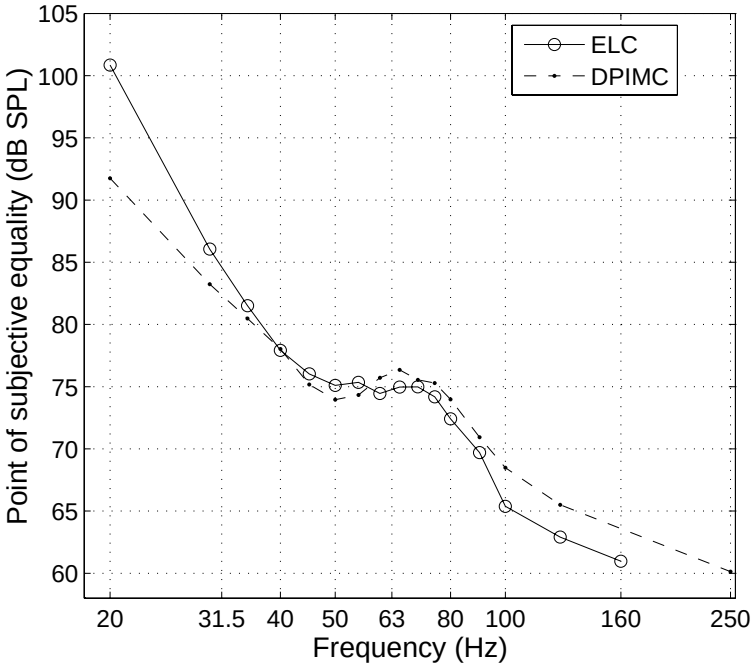

Figure 9: Example equal-loudness contour (ELC) and distortion-product-isomodulation curve (DPIMC) obtained for an individual (subject 6). The *y*-axis shows the measured levels corresponding to the ELC. The DPIMC has been vertically shifted to match the ELC. More results and details can be found in Manuscript (4).

Although the equal-loudness contours were relatively steeper, in general, a close relationship was found between perceived loudness and the objective distortion-product-isomodulation curves. This suggests that perception of loudness for pure low-frequency tones may be largely determined by the shape of the middle-ear-transfer function. Results also indicated that standardized isophon curves fail to adequately describe the observed frequency dependence of loudness perception (see e.g. ISO 226, 2003). Small resonance

features in the middle-ear-transfer function (see region between about 50 to 100 Hz in figure 9) could be identified in many cases, as the one shown in figure 9, also in the shape of individual equal-loudness contours. According to the analysis given in Manuscript (3), this region –which as suggested by Marquardt and Hensel (2008) may reflect a cochlear resonance– is involved in the observed decrease of the sharpness of tuning, described in Manuscripts (1) and (2).

In the region below about 40 Hz, the equal-loudness contours were roughly 6 dB/oct steeper than the distortion-product-isomodulation curves. The latter result allowed to improve the interpretation of the auditory filtering process inferred from the previous experiments, in that below about 40 Hz, tuning seems not solely determined by the middle-ear-transfer function, but other factors, such as inner-hair cell response patterns, may be involved.

4 General conclusions

Results from the experiments performed in this PhD project suggest that the bandwidth of the auditory filter, considering the combined effects of filtering processes at all levels in the auditory system, decreases with decreasing frequency at least down to a center frequency (CF) of 80 Hz, where it reaches a value of about 30 Hz. Below that, a worsening of tuning occurs, leading to increases in bandwidth with decreasing frequency, a tendency at odds with predictions based on simple extrapolation of previous findings. Two separate experiments applying different methods confirmed this observation. A modeling of the possible influence on tuning of the middle-ear-transfer function, which at low frequencies includes the effects of the helicotrema shunt mechanism (a cochlear process), could largely explain this phenomenon.

Besides, individual differences in bandwidth and shape of the auditory filter generally tended to increase with decreasing frequency. This seems to occur as a consequence of individual particularities observed within an irregular region of the middle-ear-transfer function. Regarding the latter transfer function, evidence was also found suggesting that perception of loudness for pure tones below 100 Hz is highly dependent on its shape, and that standardized isophon curves fail to describe this.

On the other hand, the asymmetry of tuning was found to reverse from what has been typically observed at higher frequencies, presenting steeper lower skirts than upper skirts (considering linear frequency units). This was another phenomenon associated with the middle-ear-transfer function, that occurs as a consequence of its highpass nature, and for which only few physiological evidence existed previously. As well, some evidence was found suggesting that the efficiency in the detection process may improve in the frequency range below 100 Hz. This may have been possible if listeners were able to “listen in the dips”, by taking advantage of the inherent slow amplitude fluctuations present in narrowband noise with low-frequency energy.

Analysis of the result of excluding the filtering effects of the middle-ear-transfer function from frequency selectivity, indicated that not only it is a strong factor that influences the

overall shape of the auditory filter, but that in the region below about 40 Hz, where the helicotrema shunt action is dominant, it largely defines the lower skirt of the auditory filter. Leaving aside additional effects on tuning of other potential factors, this result makes theoretical sense. This is so because very low-frequency sounds will reach the most apical regions in the cochlea, where pressure differences across the basilar membrane will be effectively shunted by the helicotrema; thus this mechanism will start becoming part of the auditory filtering process itself.

The analysis that produced the above-mentioned result, together with the fact that the psychophysical tuning curves obtained for signal frequencies of 31.5 and 40 Hz were very similar –and therefore most probably detected using the same auditory filter– led to resolve that the "bottom" auditory filter (i.e. the one with the lowest CF) is centered roughly between 40 and 50 Hz. Signals with lower frequencies than that appear to be detected via the low-frequency skirt of this auditory filter. The lowest CF is in fair agreement with the value, adopted somewhat of necessity, in the phenomenological model developed by Moore and colleagues (1997), who suggested that there are no auditory filters tuned below about 50 Hz.

5 Further work

There are different ways in which the present work can be extended. For example, the level dependence of frequency selectivity in the frequency range below 100 Hz remains to be determined. Previous studies have found little level dependence at low frequencies (see e.g Rosen and Stock, 1992; Unoki *et al.*, 2006a), although they have only measured down to a CF of 125 Hz. However, the steepness in low-frequency hearing changes noticeably with level in the frequency region below 100 Hz (see e.g. ISO 226, 2003). Besides, the fact that differences in steepness between the measured equal-loudness contours and distortion-product-isomodulation curves (both measured at relatively different levels) were largest below about 40 Hz, may suggest level dependent changes in perception in this region. Therefore, significant level dependent changes, particularly in the sharpness of the low-frequency skirt of the auditory filter, may exist. From a practical point of view, the main challenges for such an experiment are the high sound-pressure levels and large range of levels needed.

In addition, the estimates of frequency selectivity obtained in this work can be used in such a manner so as to assess the loudness produced by complex low-frequency signals, such as noise. Excitation patterns produced by specific stimuli can be obtained and used to calculate their loudness by means of a loudness model. Correlates with loudness perception for such signals can then be studied to extract conclusions. Specific relevance should be given to signals with frequency components between about 40 to 100 Hz, where larger individual differences in frequency selectivity were generally observed.

References

Aertsen, A., and Johannesma, P. (1980). "Spectro-temporal receptive fields of auditory neurons in the grassfrog: I. Characterization of tonal and natural stimuli," Biol. Cybern. 38, 223-234.

Aibara, R., Welsh, J. T., Puria, S., and Goode, R. L. (2001). "Human middle-ear sound transfer function and cochlear input impedance," Hear. Res. 152, 100-109.

Alcántara, J. I., Moore, B. C. J., and Vickers, D. A. (2000). "The relative role of beats and combination tones in determining the shapes of masking patterns at 2 kHz: I. Normal-hearing listeners," Hear. Res. 148, 63-73.

ANSI (2007). *ANSI S3.4-2007. Procedure for the computation of loudness of steady sounds* (American National Standards Institute, New York).

Arthur, R. M., Pfeiffer, R. R., and Suga, N. (1971). "Properties of 'two-tone inhibition' in primary auditory neurones," J. Physiol. 212, 593-609.

Bacon, S. P., and Moore, B. C. J. (1993). "Modulation detection interference: Some spectral effects," J. Acoust. Soc. Am. 93, 3442-3453.

Baker, R. J., and Rosen, S. (2006). "Auditory filter nonlinearity across frequency using simultaneous notched-noise masking," J. Acoust. Soc. Am. 119, 454-462.

Bilger, R. C., and Hirsh, I. J. (1956). "Masking of tones by bands of noise," J. Acoust. Soc. Am. 28, 623-630.

Blauert, J., and Laws, P. (1978). "Group delay distortions in electroacoustical systems," J. Acoust. Soc. Am. 63, 1478-1483.

Bos, C. E., and de Boer, E. (1966). "Masking and discrimination," J. Acoust. Soc. Am. 39, 708-715.

Buus, S., Schorer, E., Florentine, M., and Zwicker, E. (1986). "Decision rules in detection of simple and complex tones," J. Acoust. Soc. Am. 80, 1646-1657.

Carney, L. H., Heinz, M. G., Evilsizer, M. E., Gilkey, R. H., and Colburn, H. S. (2002). "Auditory phase opponency: A temporal model for masked detection at low frequencies," Acta Acust. - Acust. 88, 334-346.

de Boer, E. (1975). "Synthetic whole-nerve action potentials for the cat," J. Acoust. Soc. Am. 58, 1030-1045.

de Boer, E., and Bouwmeester, J. (1974). "Critical bands and sensorineural hearing loss," Audiology 13, 236-259.

de Boer, E., and Jongh, H. (1978). "On cochlear encoding: Potentialities and limitations of the reverse-correlation technique," J. Acoust. Soc. Am. 63, 115-135.

Delgutte, B. (1990). "Physiological mechanisms of psychophysical masking: Observations from auditory-nerve fibers," J. Acoust. Soc. Am. 87, 791-809.

Delgutte, B. (1996). "Physiological models for basic auditory percepts," in *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (Springer, New York), pp. 157-220.

Egan, J. P., and Hake, H. W. (1950). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. 22, 622-630.

Evans, E. F., Pratt, S. R., and Cooper, N. P. (1989). "Correspondence between behavioural and physiological frequency selectivity in the guinea pig," Br. J. Audiol. 23, 151-152.

Fidell, S., Horonjeff, R., Teffeteller, S., and Green, D. M. (1983). "Effective masking bandwidths at low frequencies," J. Acoust. Soc. Am. 73, 628-638.
Fletcher, H. (1940). "Auditory patterns," Rev. Mod. Phys. 12, 47-65.
Gässler, G. (1954). "Über die Hörschwelle für Schallereignisse mit verschieden breitem Frequenzspektrum," Acustica 4, 408-414.
Gifford, R. H., and Bacon, S. P. (2000). "Contributions of suppression and excitation to simultaneous masking: effects of signal frequency and masker-signal frequency relation," J. Acoust. Soc. Am. 107, 2188-200.
Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. 47, 103-138.
Glasberg, B. R., Moore, B. C. J., and Nimmo-Smith, I. (1984). "Comparison of auditory filter shapes derived with three different maskers," J. Acoust. Soc. Am. 75, 536-544.
Green, D. M., McKey, M. J., and Licklider, J. C. R. (1959). "Detection of a pulsed sinusoid in noise as a function of frequency," J. Acoust. Soc. Am. 31, 1446-1452.
Greenwood, D. D. (1961). "Auditory masking and the critical band," J. Acoust. Soc. Am. 33, 484-501.
Greenwood, D. D. (1971). "Aural combination tones and auditory masking," J. Acoust. Soc. Am. 50, 502-543.
Hamilton, P. M. (1957). "Noise masked thresholds as a function of tonal duration and masking noise bandwidth," J. Acoust. Soc. Am. 29, 506-511.
Hawkins, J. E., Jr., and Stevens, S. S. (1950). "The masking of pure tones and of speech by white noise," J. Acoust. Soc. Am. 22, 6-13.
Heinz, M. G., Colburn, H. S., and Carney, L. H. (2002). "Quantifying the implications of nonlinear cochlear tuning for auditory-filter estimates," J. Acoust. Soc. Am. 111, 996-1011.
Helmholtz, H. L. F. (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik (On the Sensations of Tone as a Physiological Basis for the Theory of Music)* (F.Vieweg, Braunschweig).
Houtgast, T. (1977). "Auditory-filter characteristics derived from direct-masking data and pulsation-threshold data with a rippled-noise masker," J. Acoust. Soc. Am. 62, 409-415.
Irino, T., and Patterson, R. (1997). "A time-domain, level-dependent auditory filter: The gammachirp," J. Acoust. Soc. Am. 101, 412-419.
Irino, T., and Patterson, R. D. (2001). "A compressive gammachirp auditory filter for both physiological and psychophysical data," J. Acoust. Soc. Am. 109, 2008-2022.
ISO 226 (2003). *Acoustics - normal equal-loudness contours* (International Organization for Standardization, Geneva).
Johannesma, P. (1972). "The pre-response stimulus ensemble of neurons in the cochlear nucleus," in *Symposium on Hearing Theory* (IPO, Eindhoven, The Netherlands), pp. 58-69.
Johnson-Davies, D., and Patterson, R. D. (1979). "Psychophysical tuning curves: restricting the listening band to the signal region," J. Acoust. Soc. Am. 65, 765-770.
Klippel, W. (2006). "Tutorial: Loudspeaker Nonlinearities—Causes, Parameters, Symptoms," J. Audio Eng. Soc. 54, 907-939.
Kluk, K., and Moore, B. C. J. (2004). "Factors affecting psychophysical tuning curves for normally hearing subjects," Hear. Res. 194, 118-134.

Langhans, A., and Kohlrausch, A. (1992). "Spectral integration of broadband signals in diotic and dichotic masking experiments," J. Acoust. Soc. Am. 91, 317-326.
Leshowitz, B., and Wightman, F. L. (1972). "On the importance of considering the signal's frequency spectrum: Some comments on Macmillan's 'Detection and recognition of increments and decrements in auditory intensity'," Percept. Psychophys. 12, 209-210.
Leventhall, H. G. (2004). "Low frequency noise and annoyance," Noise Health 6, 59-72.
Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (2003). "Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing," J. Acoust. Soc. Am. 113, 951-960.
Manley, G. A. (2000). "Cochlear mechanisms from a phylogenetic viewpoint," Proc. Natl. Acad. Sci. USA 97, 11736-11743.
Marquardt, T., and Hensel, J. (2008). "A lumped-element model of the apical cochlea at low frequencies," in *Concepts and Challenges in the Biophysics of Hearing*, edited by N. P. Cooper, and D. T. Kemp (World Scientific, London), pp. 337-339.
Marquardt, T., and Pedersen, C. S. (2010). "The infuence of the helicotrema on low-frequency hearing," in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York), pp. 25-36.
Marquardt, T., Hensel, J., Mrowinski, D., and Scholz, G. (2007). "Low-frequency characteristics of human and guinea pig cochleae," J. Acoust. Soc. Am. 121, 3628-3638.
Moore, B. C. J. (1978). "Psychophysical tuning curves measured in simultaneous and forward masking," J. Acoust. Soc. Am. 63, 524-532.
Moore, B. C. J. (1981). "Interactions of masker bandwidth with signal duration and delay in forward masking," J. Acoust. Soc. Am. 70, 62-68.
Moore, B. C. J. (1986). *Frequency Selectivity in Hearing* (Academic, London).
Moore, B. C. J. (2002). "Frequency resolution," in *Genetics and the Function of the Auditory System*, edited by L. Tranebjærg, J. Christensen-Dalsgaard, T. Andersen, and T. Poulsen (Holmens Trykkeri, Copenhagen, Denmark), pp. 227-271.
Moore, B. C. J. (2003). *An Introduction to the Psychology of Hearing, 5th Ed.* (Academic Press, San Diego).
Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," J. Acoust. Soc. Am. 74, 750-753.
Moore, B. C. J., and Glasberg, B. R. (1987). "Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes," J. Acoust. Soc. Am. 82, 69-79.
Moore, B. C. J., and Sek, A. (1995). "Auditory filtering and the critical bandwidth at low frequencies," in *Advances in Hearing Research*, edited by G. A. Manley, G. M. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 425-436.
Moore, B. C. J., and Vickers, D. A. (1997). "The role of spread of excitation and suppression in simultaneous masking," J. Acoust. Soc. Am. 102, 2284-2290.
Moore, B. C. J., Alcántara, J. I., and Dau, T. (1998). "Masking patterns for sinusoidal and narrowband noise maskers," J. Acoust. Soc. Am. 104, 1023-1038.
Moore, B. C. J., Glasberg, B. R., and Baer, T. (1997). "A model for the prediction of thresholds, loudness and partial loudness," J. Audio Eng. Soc. 45, 224-240.

Moore, B. C. J., Glasberg, B. R., and Roberts, B. (1984). "Refining the measurement of psychophysical tuning curves," J. Acoust. Soc. Am. 76, 1057-1066.
Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1990). "Auditory filter shapes at low center frequencies," J. Acoust. Soc. Am. 88, 132-140.
Narayan, S. S., Temchin, A. N., and Ruggero, M. A. (1998). "Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae," Science 282, 1882 - 1884.
Niemiec, A. J., and Yost, W. A. (1992). "Behavioral measures of frequency selectivity in the chinchilla," J. Acoust. Soc. Am. 92, 2636-2649.
O'Loughlin, B. J., and Moore, B. C. J. (1981). "Improving psychoacoustical tuning curves," Hear. Res. 5, 343-346.
Palmer, A. R. (1987). "Physiology of the cochlear nerve and cochlear nucleus," in *Hearing*, edited by M. P. Haggard, and E. F. Evans (Churchill Livingstone, Edinburgh), pp. 838-855.
Patterson, R., Nimmo-Smith, I., Holdsworth, J., and Rice, P. (1987). An efficient auditory filter bank based on the gammatone function. Paper presented at a meeting of the IOC Speech Group on Auditory Modelling at RSRE. December 14-15.
Patterson, R. D. (1974). "Auditory filter shape," J. Acoust. Soc. Am. 55, 802-809.
Patterson, R. D. (1976). "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. 59, 640-654.
Patterson, R. D. (1987). "A pulse ribbon model of monaural phase perception," J. Acoust. Soc. Am. 82, 1560-1586.
Patterson, R. D., and Henning, G. B. (1977). "Stimulus variability and auditory filter shape," J. Acoust. Soc. Am. 62, 649-664.
Patterson, R. D., and Moore, B. C. J. (1986). "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 123-177.
Patterson, R. D., and Nimmo-Smith, I. (1980). "Off-frequency listening and auditory filter asymmetry," J. Acoust. Soc. Am. 67, 229-245.
Patterson, R. D., Unoki, M., and Irino, T. (2003). "Extending the domain of center frequencies for the compressive gammachirp auditory filter," J. Acoust. Soc. Am. 114, 1529-1542.
Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). "The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold," J. Acoust. Soc. Am. 72, 1788-1803.
Pickles, J. O. (1984). "Frequency threshold curves and simultaneous masking functions in single fibers of the guinea pig auditory nerve," Hear. Res. 14, 245-256.
Puria, S., Rosowski, J. J., and Peake, W. T. (1997). "Sound-pressure measurements in the cochlear vestibule of human-cadaver ears," J. Acoust. Soc. Am. 101, 2754-2770.
Rabinowitz, W. M., Bilger, R. C., Trahiotis, C., and Neutzel, J. (1980). "Two-tone masking in normal hearing listeners," J. Acoust. Soc. Am. 68, 1096-1106.
Rosen, S., and Stock, D. (1992). "Auditory filter bandwidths as a function of level at low frequencies (125 Hz-1 kHz)," J. Acoust. Soc. Am. 92, 773-781.
Rosen, S., Baker, R. J., and Darling, A. (1998). "Auditory filter nonlinearity at 2 kHz in normal hearing listeners," J. Acoust. Soc. Am. 103, 2539-2550.

Ruggero, M. A., Robles, L., and Rich, N. C. (1992). "Two-tone suppression in the basilar membrane of the cochlea: Mechanical basis of auditory-nerve rate suppression," J. Neurophysiol. 68, 1087-1099.
Sachs, M. B., and Kiang, N. Y. S. (1968). "Two-tone inhibition in auditory nerve fibers," J. Acoust. Soc. Am. 43, 1120-1128.
Schaeffer, T. H., Gales, R. S., Shewmaker, C. A., and Thompson, P. O. (1950). "The frequency selectivity of the ear as determined by masking experiments," J. Acoust. Soc. Am. 22, 490-496.
Scharf, B. (1970). "Critical bands," in *Foundations of Modern Auditory Theory*, edited by J. V. Tobias (Academic Press, New York), pp. 157-202.
Shailer, M. J., Moore, B. C. J., Glasberg, B. R., Watson, N., and Harris, S. (1990). "Auditory filter shapes at 8 and 10 kHz," J. Acoust. Soc. Am. 88, 141-148.
Shera, C. A., Guinan, J. J., Jr., and Oxenham, A. J. (2002). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proc. Natl. Acad. Sci. USA 99, 3318-3323.
Small, A. M. (1959). "Pure-tone masking," J. Acoust. Soc. Am. 31, 1619-1625.
Spiegel, M. F. (1981). "Thresholds for tones in maskers of various bandwidths and for signals of various bandwidths as a function of signal frequency," J. Acoust. Soc. Am. 69, 791-795.
Terry, M., and Moore, B. C. J. (1977). "'Suppression' effects in forward masking," J. Acoust. Soc. Am. 62, 781-784.
Unoki, M., Kazuhito, I., Ishimoto, Y., and Tan, C. T. (2006a). "Estimate of auditory filter shape using notched-noise masking for various signal frequencies," Acoust. Sci. Tech. 27, 1-11.
Unoki, M., Irino, T., Glasberg, B. R., Moore, B. C. J., and Patterson, R. D. (2006b). "Comparison of the roex and gammachirp filters as representations of the auditory filter," J. Acoust. Soc. Am. 120, 1474-1492.
Vogten, L. L. (1978). "Low-level pure-tone masking: a comparison of "tuning curves" obtained with simultaneous and forward masking," J. Acoust. Soc. Am. 63, 1520-1527.
Vogten, L. L. M. (1974). "Pure-tone masking: A new result from a new method," in *Facts and Models in Hearing*, edited by E. Zwicker, and E. Terhardt (Springer-Verlag, Berlin), pp. 142-155.
Wegel, R. L., and Lane, C. E. (1924). "The auditory masking of one sound by another and its probable relation to the dynamics of the inner ear," Phys. Rev. 23, 266-285.
Yost, W. A., Sheft, S., and Opie, J. (1989). "Modulation interference in detection and discrimination of amplitude modulation," J. Acoust. Soc. Am. 86, 2138-2147.
Zwicker, E. (1952). "Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones (The limits of audibility of amplitude modulation and frequency modulation of a pure tone)," Acustica 2, 125-133.
Zwicker, E. (1954). "Die Verdeckung von Schmalbandgeräuschen durch Sinustöne," Acustica 4, 415-420.
Zwicker, E. (1974). "On the psychophysical equivalent of tuning curves," in *Facts and Models in Hearing*, edited by E. Zwicker, and E. Terhardt (Springer-Verlag, Berlin), pp. 132-140.
Zwicker, E., and Scharf, B. (1965). "A model of loudness summation," Psych. Rev. 72, 3-26.

Zwicker, E., and Schutte, H. (1973). "On the time pattern of the threshold of tone impulses masked by narrow band noise," Acustica 29, 343-347.
Zwicker, E., Flottorp, G., and Stevens, S. S. (1957). "Critical bandwidth in loudness summation," J. Acoust. Soc. Am. 29, 548-557.

Frequency selectivity for frequencies below 100 Hz: comparisons with mid-frequencies

Carlos A. Jurado[a)]
Section of Acoustics, Department of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7-B5, 9220 Aalborg Ø, Denmark

Brian C.J. Moore
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

a) email: cjo@es.aau.dk

Abstract

Auditory filter shapes were derived for signal frequencies (f_S) between 50 and 1000 Hz, using the notched-noise method. The masker spectrum level (N_0) was 50 dB (re 20 μPa). For f_S = 63 and 50 Hz, measurements were also made with N_0 = 62 dB for the lower band. The data were fitted using a rounded-exponential filter model, with special consideration of the filtering effects of the middle-ear transfer function (METF) at low frequencies. The results showed: (1) For very low values of f_S, the lower skirts of the filters were only well defined when N_0 = 62 dB for the lower band; (2) The sharpness of both sides of the filters decreased with decreasing f_S; (3) The dynamic range of the filters decreased with decreasing f_S; (4) The equivalent rectangular bandwidth of the filters decreased with decreasing f_S down to f_S = 80 Hz, but increased for f_S below that; (5) The assumed METF, which includes the shunt effect of the helicotrema for frequencies below 50 Hz, increasingly influenced the low-frequency skirt of the filters as f_S decreased; (6) Detection efficiency worsened with decreasing f_S for f_S between 100 and 500 Hz, but improved slightly below that.

PACS number(s): 43.66.Ba, 43.66.Dc, 43.66.Cb

I. INTRODUCTION

Many environmental sounds contain substantial energy at low frequencies. Such sounds, either of community or industrial origin, can propagate over long distances and cause problems with annoyance and masking (Leventhall, 2004). While the audibility and loudness of low-frequency pure tones can be predicted from absolute threshold curves and equal-loudness-level contours (ISO 226, 2003), respectively, the audibility and loudness of complex sounds with substantial low-frequency energy cannot be predicted in this way. Instead, such predictions need to be based on auditory models, which in turn depend on information about the characteristics of the auditory filters (see e.g. Moore *et al.*, 1997).

Although auditory filter characteristics have been studied over most of the audible frequency range (Glasberg and Moore, 1990; Shera *et al.*, 2002; Baker and Rosen, 2006; Unoki *et al.*, 2006b; Rosen and Stock, 1992), not much work has been dedicated to

the low-frequency region, below 250 Hz. There is some evidence indicating that the equivalent rectangular bandwidth (ERB) of the human auditory filter decreases with decreasing center frequency (CF) for CFs down to about 100 Hz (Moore *et al.*, 1990; Peters and Moore, 1992; Moore and Sek, 1995; Rosen and Stock, 1992). Based on data for center frequencies above 100 Hz, Glasberg and Moore (1990) proposed an equation relating the average ERB of the auditory filters for young listeners with normal hearing, denoted ERBN, to CF:

$$ERB_N = 24.7 \times (4.37CF/1000+1) \qquad (1)$$

This equation predicts that the value of ERB_N should flatten off for CFs below 100 Hz, reaching a value of about 30 Hz for CF = 50 Hz and about 26 Hz for CFs below 20 Hz. However, it is unrealistic to expect there will be auditory filters centered across the whole of this range. Since the auditory filters are assumed to depend mainly on cochlear processes, there will be some limiting low frequency at which the apical end of the cochlea is reached. Moore *et al*. (1997) assumed that there were no auditory filters with CFs below 50 Hz. They argued that sounds with frequencies below 50 Hz were detected via the low-frequency skirt of the auditory filters with the lowest CFs. However, the CF of the lowest auditory filter remains uncertain, as do the values of ERB_N for CFs below 100 Hz. One aim of the present study was to provide more information about the shape of the auditory filter and the values of ERB_N for very low CFs. This information is important both for improving our understanding of frequency selectivity at low frequencies and for further development of models that can be used for assessment and prediction of problems caused by low-frequency sounds.

Methods for deriving the shape of the auditory filter are typically based on the power-spectrum model of masking[1] (Fletcher, 1940; Patterson and Moore, 1986). According to this model, the threshold for detecting a tone in a noise is influenced both by the shape of the auditory filter and by "detection efficiency", which is related to the signal-to-masker ratio at the output of the auditory filter required to reach the detection threshold, denoted *K*; a high value of *K* indicates poor efficiency. The power-spectrum model makes it possible to separate these two factors. Previous studies have shown that *K* tends to increase with decreasing CF, reaching values above 0 dB for CFs around 100 Hz (Moore *et al.*, 1990; Unoki *et al.*, 2006a). One factor that may be partially responsible for the increase in *K* is that the number of stimulus cycles for a signal of fixed duration decreases with decreasing frequency.

Inherent amplitude fluctuations in the noise maskers may also influence *K* at low frequencies, because of the reduced bandwidth of the signals passing through the auditory filter, which leads to slower fluctuations. Slower fluctuations may impair performance based on increases in intensity produced by adding the signal to the masker (Bos and de Boer, 1966). However, if the masker fluctuations become very slow, listeners may be able to detect the signal in minima of the masker envelope, which could lead to improved performance (Zwicker and Schutte, 1973; Moore and Glasberg, 1987a). It is therefore not clear how *K* might vary with frequency at very low CFs. Another objective of this study was to estimate values of *K* for very low CFs.

Psychophysical measures of frequency selectivity are influenced by processes at many levels in the auditory system. Conceptually, one can distinguish between fixed filtering processes, arising from the transfer of sound through the outer and middle ear, and frequency-selective

processes occurring mainly in the cochlea or later (Glasberg and Moore, 1990). The auditory filter shape is usually assumed to be related to the second of these two. Data from masking experiments can be analyzed, based on the power-spectrum model, by treating these two aspects separately. This is done by modifying the spectra of the stimuli using fixed filters to simulate the effect of the outer/middle ear, and using these modified spectra as input to the filter model. For low frequencies, the outer ear has little effect (Shaw, 1974), but the middle-ear transfer function (METF) can have a strong effect (Rosowski, 1996; Puria *et al.*, 1997). The METF is often defined as the ratio of the sound pressure difference across the basilar membrane to the sound pressure at the tympanic membrane (Dallos, 1973; Marquardt *et al.*, 2007; Marquardt and Hensel, 2008; Marquardt and Pedersen, 2010). If the METF is taken into account, the lower skirt of the auditory filter at low CFs is relatively shallow. However, if the effect of the METF is ignored in the analysis (so that its effects are included in the fitted filter shape), the lower skirt of the derived filter tends to be steeper (Baker and Rosen, 2006; Unoki et al., 2006a).

Below about 50 Hz, the METF increases markedly in steepness, which can be attributed to the shunt action of the helicotrema (Cheatham and Dallos, 2001; Marquardt et al., 2007; Dallos, 1970; Marquardt and Hensel, 2008). The helicotrema reduces the pressure difference between scala vestibuli and scala tympani, and this leads to a decrease in sensitivity to low-frequency sound. Therefore, the influence of the METF on measures of frequency selectivity should be particularly pronounced in this region. Direct measurements of tuning at apical regions of guinea pig cochleae indicate that the effect of the METF can become so dominant that the lower side of the tuning curve is steeper than upper side, the opposite of the situation for higher CFs (Cheatham and Dallos, 2001).

An issue that remains unresolved is whether, at low frequencies, the METF should be considered as part of frequency selectivity or not. As described above, in previous work the effect of the METF has been modeled as a fixed filter applied prior to auditory filtering (Glasberg and Moore, 1990). However, as the displacement of the middle-ear ossicle chain is stiffness dominated below about 500 Hz (Puria *et al.*, 1997; Marquardt *et al.*, 2007), the further increase in the steepness of the METF below about 50 Hz is mainly a consequence of the shunt action of the helicotrema, which is equivalent to a decrease in cochlear impedance at the apex (Schick, 1994; Marquardt *et al.*, 2007). This is a cochlear mechanism and is not part of the function of the middle ear per se (the ossicular chain). Thus, the term METF could be considered as misleading, as the METF includes a cochlear component. Nevertheless, it seems reasonable to treat the METF as a fixed effect that occurs prior to auditory filtering, since the effect of the METF appears to be independent of level and CF. The effect of including the METF as part of auditory filtering was considered in the analysis described in the present paper, with focus on its effect for very low CFs. This was done to estimate the importance of the METF in determining the tuning of the system as a whole.

In this study, auditory-filter characteristics for normal-hearing subjects were estimated for CFs in the range 50-1000 Hz, using the notched-noise method (de Boer and Bouwmeester, 1974; Patterson, 1976; Patterson and Nimmo-Smith, 1980). The objective was to estimate the auditory-filter shape and bandwidth at low CFs, to estimate detection efficiency at low CFs, and to evaluate the influence of the METF

on estimates of frequency selectivity at low CFs. The range of CFs was chosen to include CFs that have previously been studied, to allow comparison with earlier work and to facilitate across-CF comparisons. For CFs of 63 and 50 Hz, a new methodological approach was used where the lower of the two masking bands was increased in level relative to the upper band, to compensate for the expected effects of METF attenuation, thus allowing more accurate estimates of the shape of the lower skirt of the auditory filter.

II. METHOD

There has been some controversy in the literature about whether it is more appropriate to estimate auditory filter shapes using a fixed masker level or a fixed signal level (Rosen *et al.*, 1998; Glasberg and Moore, 2000; Baker and Rosen, 2006). This can be important when the auditory filter shape is strongly level dependent. However, for low CFs, the shape of the auditory filter changes only slightly with level (Rosen and Stock, 1992; Unoki *et al.*, 2006b), so similar estimates of auditory filter shape should be obtained using the two methods. Here, we chose to use a fixed masker level. In other studies, we have gathered data using a fixed signal level (Jurado et al., 2010a).

A. Notched-noise spectral configurations and signal frequencies

Masked thresholds were obtained for pure-tone signals presented in a spectral notch between two bands of noise, each with a width of $0.4f_S$, where f_S is the signal frequency. Eight values of f_S were used: 50, 63, 80, 100, 125, 250, 500, and 1000 Hz. The notch was placed both symmetrically and asymmetrically around f_S. In the symmetrical case, the normalized frequency separation, $\Delta f/f_S$, between the spectral edges of the noise masker closest to f_S and f_S were 0, 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6. Two asymmetric band-placement cases were used. In the 'upper+' case, the values of $\Delta f/f_S$ for the lower band were 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6 while the corresponding values for the upper band were 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8. In the 'lower+' case, the values of $\Delta f/f_S$ for the upper band were 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6 while the corresponding values for the lower band were 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8. In cases where frequency components of the lower band would have fallen below 0 Hz, the bandwidth was reduced accordingly.

B. Threshold procedure

A three-alternative forced-choice task was used. Responses were made via a box with three buttons. Feedback was provided after each response. A simple 1-up 1-down rule was used for the first four trials, to achieve rapid convergence to the region of masked threshold. After that, a 3-down 1-up adaptive procedure tracking the 79% correct point on the psychometric function was used to estimate masked threshold (Levitt, 1971). The starting level of the signal was set on the basis of pilot trials so that the signal was clearly audible. The step size started at 8 dB, was reduced to 4 dB after two turnpoints, and was further reduced to 2 dB after two further turnpoints, where it remained.

Twelve turnpoints were obtained and threshold was taken as the average of the signal levels at the last eight turnpoints (i.e. all turnpoints obtained using the 2-dB step size). If the standard deviation (SD) of the levels at the last eight turnpoints exceeded 4 dB, the estimate for that single threshold measurement (i.e. run) was discarded and a new estimate was obtained. In a session including 19 notched-noise conditions (12 asymmetric and 7 symmetric conditions), the number of discarded runs ranged from 0

to 3, depending on the subject. For each condition, the average of two threshold estimates was taken as masked threshold, except when the two estimates differed by more than 3 dB, in which case a third estimate was obtained and all three were averaged. Absolute thresholds were obtained in essentially the same manner.

C. Subjects

A total of 11 subjects participated. They were divided into two groups. Group 1 was tested using a wide range of signal frequencies while Group 2 was tested using more closely spaced signal frequencies in the low-frequency range. Group 1 consisted of two males and three females, aged 25 to 29 years. They were tested using f_S = 50, 63, 125, 250, 500, and 1000 Hz. Group 2 consisted of three males and three females, aged 21 to 25 years. They were tested using f_S = 50, 63, 80, 100, and 125 Hz. Four of these subjects were tested using all values of f_S, while two subjects were tested only for f_S = 50 and 63 Hz. Overall, 11 subjects were tested for f_S = 50 and 63 Hz, nine for f_S = 125 Hz, five for f_S = 250, 500, and 1000 Hz, and four for f_S = 80 and 100 Hz. Subjects were selected to have audiometric thresholds <15 dB HL at the standard audiometric frequencies, and differences in threshold between ears of up to 10 dB were accepted. All subjects were given 2 hours of practice. Breaks were given regularly after 12-15 minutes of testing (corresponding to about four threshold measurements), in order to maintain concentration.

D. Stimuli and equipment

Notched-noise stimuli were created by filtering a wideband noise signal (>30 s long) with cascaded bandstop, highpass, and lowpass digital 200th-order infinite impulse response filters (Chebychev Type II), formed by cascading 100 second-order sections, and designed using the filter design toolbox in Matlab. The filter slopes were extremely sharp; the response changed by 80 dB for a step in frequency of 1.7 Hz. Noise bursts that were 600-ms long, including 50-ms linear ramps at their start and end, were obtained by taking a random sample from the 30-s noise buffer. The three noise bursts within a trial were separated by silent intervals of 500 ms. To minimize the effects of fluctuations in overall amplitude, the three noise bursts within a given trial were identical, based on the same randomly chosen noise sample. The noise sample used was randomly selected for each trial.

Tests using f_S = 50 Hz showed that masked thresholds obtained with these stimuli did not differ from those obtained using continuous (running) noise. Thus we believe that the use of three identical noise bursts for each trial did not influence the outcome. The tone signal was 500-ms long, including 25-ms ramps at the start and end. Spectral analyses of the tone signal and noise maskers indicated that the spectral notches in the maskers were well represented and that spectral splatter produced by the gating of the signal would have been masked by the noise.

The tone was presented randomly in one of the three noise bursts in a trial, starting 50 ms after the start of the noise burst and finishing 50 ms earlier. Signals were sent through the ADAT optical outputs of an RME DIGI 96 PC soundcard to an RME ADI-8 D/A converter, using a sample rate of 48 kHz and 24-bit resolution. The signals were passively attenuated, and presented in a sound-isolated listening room diotically through Beyerdynamic DT-990 headphones driven by a Pioneer A-616 amplifier.

To check on the levels of harmonic distortion produced by the headphones, a

sinusoidal signal with a level of 90 dB SPL (the maximum allowed level) was generated and the output was measured using an artificial ear (Bruel and Kjær type 4153). The levels of the 2nd, 3rd , and 4th harmonics were, respectively, at least 48, 49, and 65 dB lower than the level of the sinewave for all values of f_S used. For input frequencies below 50 Hz, the relative levels of the harmonics increased slightly, but, even for an input signal at 16 Hz, the levels of all harmonics were at least 28 dB below the level of the input sinewave. At the levels actually used, which were typically at least 20 dB lower than 90 dB SPL, the relative levels of the harmonics would have been even lower, so they would almost certainly have been inaudible.

E. Compensation for the headphone response

The headphone frequency response was measured using a maximum-length sequence system analyzer (DRA laboratories, Rife and Vanderkooy, 1989). The response was measured while the subject was wearing a miniature microphone fitted, using a foam earplug, flush with the ear canal entrance, as recommended by Møller et al. (1995). The response was measured three times for each ear, the headphone being repositioned between measurements. The average of the three measurements was taken as the final response. A digital filter was then designed to compensate for the headphone response so as to provide an overall flat frequency response down to about 16 Hz. The inverse filter was designed using a frequency-sampling based method (Jackson, 1996). This was done separately for each ear of each subject. Examples of responses measured at the ear and responses after individual calibration are given in Fig. 1. The compensated responses are flat (within 2 dB from 0 dB reference) for frequencies above 16 Hz. A gain limitation was applied at the lowest frequencies (below about 16 Hz) to avoid distortion.

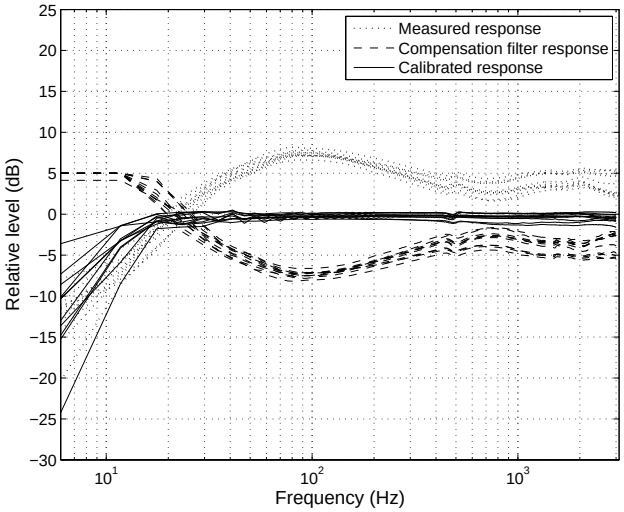

FIG. 1. Individual calibrations of the Beyerdynamic DT-990 headphone (left ear, 11 subjects). The dotted lines show the average of three responses measured for each left ear. The dashed lines show the frequency responses of the individual compensation filters, which are based on the averaged left-ear frequency responses. The solid lines show individually calibrated frequency responses obtained after filtering single left-ear measured impulse responses (measured on the corresponding subject) with each compensation filter.

F. Masker levels

For signal frequencies in the range 80 to 1000 Hz, the at-ear masker spectrum level (N_0) was 50 dB (re 20 μPa) for both bands. This was also the case for signal frequencies of 63 and 50 Hz for the subjects of Group 1. For Group 2, for f_S = 63 and 50 Hz, the value of N_0 for the upper band was kept at 50 dB, but N_0 for the lower band level was increased to 62 dB. This was done to compensate for the expected effects of attenuation produced by the METF, which increases in steepness below about 50 Hz.

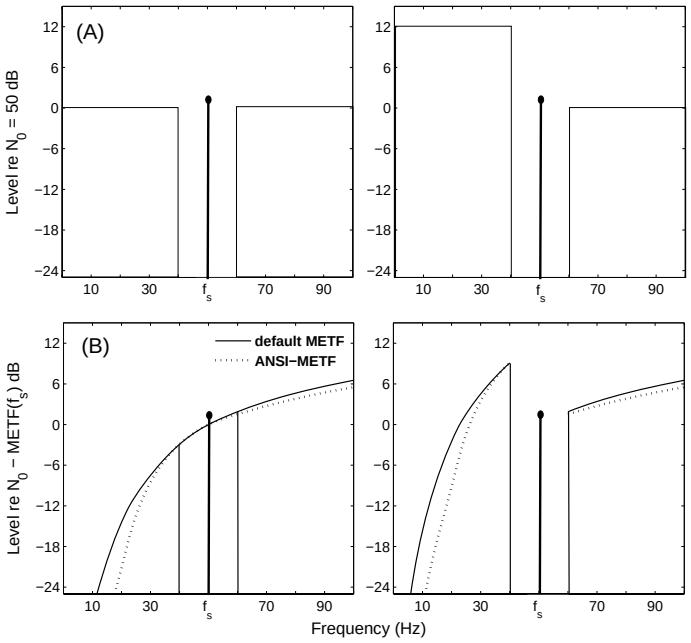

FIG. 2. (A) The spectra of notched noises without (left) and with (right) LF emphasis, for f_S = 50 Hz. (B) Corresponding spectra after filtering with the METF. The solid line indicates the characteristics of the default METF used here. The dotted line shows the ANSI-METF specified in ANSI (2007).

The increase in level was intended to ensure that noise components falling below the CF of the auditory filter would significantly influence the filter output, while avoiding distortion or excessive loudness, thus making it easier to measure the characteristics of the low-frequency skirt of the auditory filter. In what follows, conditions with the increased level of the lower band will be described as having 'LF emphasis.' Figure 2(A) compares the spectra of stimuli without (left) and with (right) LF emphasis for f_S = 50 Hz. Figure 2(B) shows the expected spectra after allowing for the effects of the "default" METF (solid line, see below for details).

III. MODELING OF THE RESULTS

A. Auditory-filter model

The auditory filter was modeled using the roex(p, r) filter shape described by Patterson et al. (1982). The relative power response of each side of the filter, W(g), is defined as:

$$W(g) = (1-r)(1+pg)\exp(-pg) + r \qquad (2)$$

where g is the deviation from the CF of the filter divided by the CF, p is a parameter which defines both the tip bandwidth and the sharpness of the slope of the filter, and r is a parameter which determines the dynamic range of the filter. The parameter p was allowed to differ for the lower and upper skirts of the filter; the values are denoted p_l and p_u. The value of r was set to the same value for the two skirts. The amount of filter asymmetry was quantified using the asymmetry index: $A = p_u/p_l$ (Glasberg *et al.*, 1984).

B. Allowing for the middle-ear transfer function

The low-frequency characteristics of the METF were assumed to have the form suggested by the work of Marquardt *et al.* (2007), Marquardt and Hensel (2008), and Marquardt and Pedersen (2010), except that no irregularity in the region around 50 Hz was assumed. The slope was approximately 12 dB/oct for frequencies below 50 Hz and 6 dB/oct for frequencies from 50 up to 150 Hz. The curve therefore includes the shunt effect of the helicotrema which was assumed to be dominant for frequencies below 50 Hz (Hensel *et al.*, 2007; Marquardt *et al.*, 2007). The form of the METF for frequencies above 150 Hz was made consistent with the METF described in the

ANSI standard for calculation of loudness (ANSI, 2007). The whole function was scaled so that the gain at 1000 Hz was 0 dB. For comparison purposes the data were also analyzed using the METF described in the ANSI S3.4 standard. When used, this is referred to as the ANSI-METF. Both METFs show an increase in steepness below about 50 Hz, but this effect is greater for the ANSI-METF than for the 'default' METF used here.

Since substantial individual differences in METFs can be expected (Aibara *et al.*, 2001; Puria *et al.*, 1997), the assumed METF at best represents the average across a group of listeners. Furthermore, the exact location of the helicotrema shunt point can be rather individual (Pedersen and Marquardt, 2009; Marquardt and Pedersen, 2010; Jurado *et al.*, 2010b). Therefore, the analyses are based on the data averaged across subjects.

C. Fitting procedure

The auditory filter model and METF were used to fit the data using the assumptions of the power-spectrum model of masking (Fletcher, 1940; Patterson et al., 1982; Patterson and Moore, 1986). The fitting procedure was similar to the one described by Glasberg and Moore (1990). For each f_S (50, 63, 80, 100, 125, 250, 500 and 1000 Hz), the three parameters described earlier were used to characterize the filter, namely p_l, p_u, and r. An additional parameter was the measure of detection efficiency, K, defined as the signal-to-noise ratio at the output of the filter required for threshold. For a given f_S, a starting set of parameters was assumed. For each notch width, it was assumed that the subject made use of the auditory filter with the highest signal-to-masker ratio at its output. The CF of this filter, which was generally not exactly equal to f_S, was first determined. In other words, off-frequency listening was taken into account. The shifted filter was used to predict the signal threshold for that notch width. This was done for each notch width in turn. The maximum allowed shift of the CF of the filter was set to $0.3f_S$. However, this limit was seldom reached, and typically the shifts were small. When the maximum shift was increased to $0.6f_S$, the limit was never reached, and typically the fits were similar to those obtained with the $0.3\ f_S$ limit.

The value of K was chosen so that the mean of the predicted thresholds was equal to the mean of the obtained thresholds. The root-mean-square deviation between the predicted and obtained thresholds was determined; this is denoted root-mean-square error (RMSe). The values of the parameters were iteratively adjusted, using the Levenberg-Marquardt algorithm (Marquardt, 1963), to minimize the RMSe. To ensure that procedure hadn't reached a local minimum in the parameter space rather than a global minimum, the procedure was re-initialized with a different set of initial parameter values (either related to the previous iteration outcome or to a fully different set of starting parameters). Re-initialization was done several times until the same (lowest) minimum was consistently found.

In the main analysis it was assumed that the auditory filter was preceeded by a METF with characteristics as described above. This analysis is referred to as "METF separate". However, the frequency region in the METF below 50 Hz, where the helicotrema shunt effect is assumed to be dominant, was considered to be special, and estimates of auditory filter shape were obtained both including and excluding this region of the METF as part of the auditory filter. Analysis based on the at-ear levels of the signal and masker, without allowance for the METF, was also done, as a way of

characterizing the frequency selectivity of the system as a whole, and to allow comparison with previous work (e.g. that of Rosen and Stock, 1992). This will be refered to as the "System-as-a-whole" analysis.

IV. RESULTS AND DISCUSSION

A. General observations

Generally, the pattern of the results was similar across subjects. The SD of the masked thresholds was typically about 1.3 dB for a notch width of 0, but increased as the notch width increased (to about 3.5 dB for a symmetric notch width of 0.6), presumably reflecting individual differences in sharpness of tuning. Figure 3 shows the mean data and fits from the model obtained in the METF-separate analysis for f_S in the range 80-1000 Hz. Mean data and fits from the model obtained at 63 and 50 Hz with and without LF-emphasis are shown in Fig. 4(A) and (B). The range of masked thresholds for a given f_S showed a clear trend to decrease with decreasing f_S. For the largest notch widths, mean masked thresholds were always 9 dB or more above absolute threshold (horizontal dashed-dotted lines). For the individual data, the masked threshold was always 5 dB or more above the absolute threshold. Thus, the decreasing range of masked thresholds is probably not solely a consequence of the moderate masker level used but probably also reflects a broadening of the filters (relative to CF) as the CF decreased.

For f_S = 1000 and 500 Hz, the data for the 'upper+' conditions (right-pointing triangles) fell above the data for the 'lower+' conditions (left-pointing triangles). This is consistent with earlier work for similar masker levels (Moore and Glasberg, 1987b; Glasberg and Moore, 2000; Baker and Rosen, 2006), and it has been interpreted as indicating that the upper skirt of the auditory filter is steeper than the lower skirt. However, this asymmetry was small for f_S from 250 to 80 Hz, and, when no LF-emphasis was applied, it was even slightly reversed for f_S = 63 and 50 Hz (Fig. 4(A)). This means that the masked threshold became increasingly dominated by the upper noise band as f_S was decreased. This was expected, as the lower band would have been progressively more attenuated by the METF as the CF of the filter decreased.

When LF-emphasis was applied (Fig. 4(B)), this would have markedly increased the contribution of the lower noise band to the masked threshold. Consistent with this, the data for the 'upper+' conditions (right-pointing triangles) fell above the data for the 'lower+' conditions (left-pointing triangles).

B. Derived filter parameters

Derived parameters are shown in Tables I and II, for both the METF-separate and the System-as-a-whole analyses. For f_S = 50 and 63 Hz (Table II), where data were obtained both without and with LF-emphasis, the filter model was fitted separately for these two sets of conditions, and to the combined data for the two conditions. When the combined data were fitted, it was assumed that the filter was level independent; the filter parameters were not made dependent on the signal or noise level.

Given that the filter was characterized by only three parameters, the model accounted for the data reasonably well; RMSe values were generally about 1 dB or less. There were no systematic deviations of the data from the predicted values except perhaps for the condition with LF emphasis for f_S = 50 Hz; when the normalized deviation of the nearer noise spectral edge from f_S was 0.2, the data point for the 'upper+' case fell slightly above the data point for the

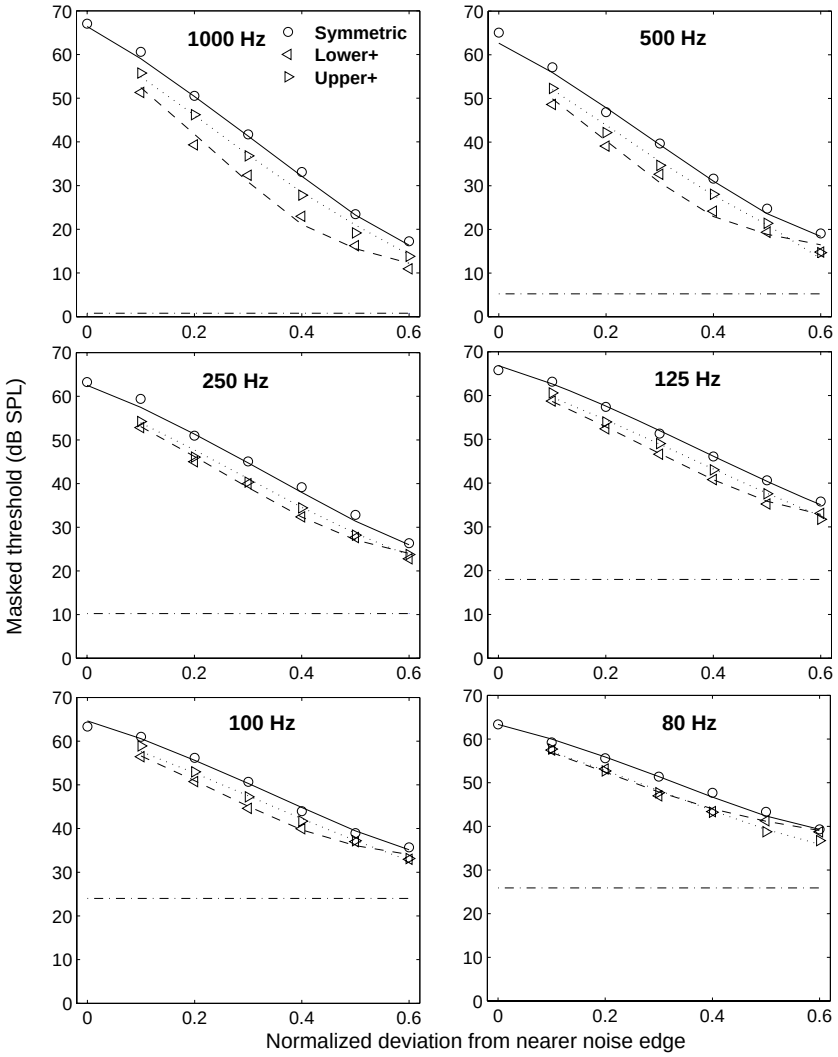

FIG. 3. Mean masked thresholds and fits from the model for f_S = 1000, 500, 250, 125, 100, and 80 Hz. Thresholds for symmetric-notch conditions are denoted by circles and corresponding fits by solid lines. Thresholds for 'lower+' conditions are denoted by left-pointing triangles and corresponding fits by dashed lines. Thresholds for 'upper+' conditions are denoted by right-pointing triangles and corresponding fits by dotted lines. Mean absolute thresholds are shown by horizontal dashed-dotted lines.

symmetric-notch case, an effect which was not predicted by the model. The reason why this occurred is not clear. It may have been due to random errors of measurement (the mean measured thresholds may sometimes be higher or lower than the "true" values).

1. Values of p

Consider first the values of the filter slope parameters, p_u and p_l, shown in Tables I and II (see LF-emphasis conditions only for f_s = 50 and 63 Hz) and plotted in Fig. 5(A) for the METF-separate analysis. These show a trend to decrease as f_s decreased from 1000 to 50 Hz, indicating a progressive broadening of the auditory filter when bandwidth is expressed relative to CF. The values of the asymmetry parameter, *A*, are plotted in Fig. 5(B) (solid line). For most values of f_s, p_u was greater than p_l. This

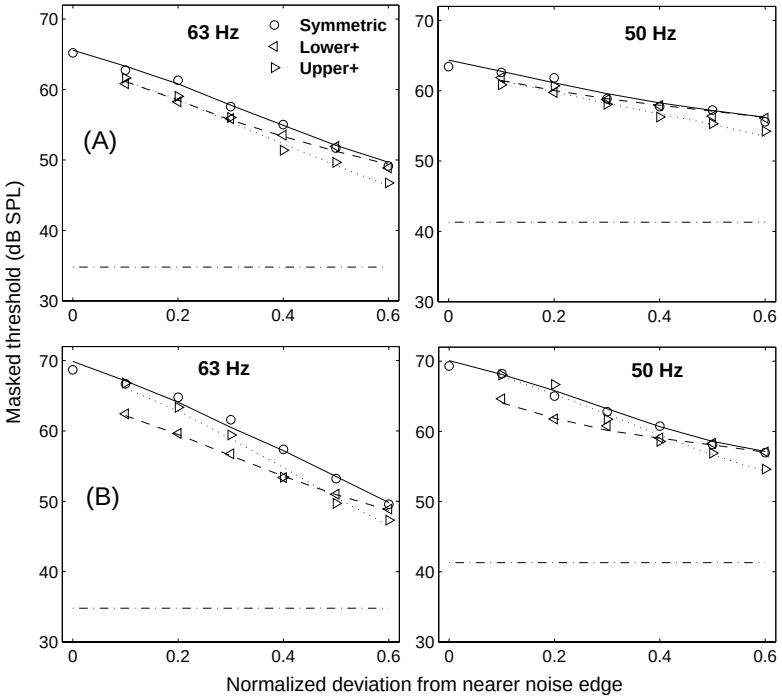

FIG. 4. As Fig. 3, but for f_S = 63 and 50 Hz. (A) cases without LF emphasis; (B) cases with LF emphasis.

became especially marked at 50 Hz, where p_l had a low value, probably because most of the observed tuning was accounted for by the steep slope of the assumed METF at very low frequencies; for a qualitatively similar effect, see Baker and Rosen (2006).

For comparison, the asymmetry values for the System-as-a-whole analysis are shown as the dashed line in Fig. 5(B). For the higher values of f_s, these were similar to those obtained for the METF-separate analysis, and are consistent with the pattern of the data discussed above (left-pointing triangles mostly below right-pointing triangles in Fig. 3). However, as f_s decreased, differences between the two analysis cases became increasingly marked. In particular, below 100 Hz, p_l became larger than p_u for the System-as-a-whole analysis, a clear reversal in asymmetry from the pattern observed at higher values of f_s. Such 'reverse' asymmetry (p_l greater than p_u) is consistent with the physiological measures of tuning reported by Cheatham and Dallos (2001), which are comparable to the System-as-a-whole analysis described here. The reversal appears to be a consequence of the highpass nature of the METF. It is also consistent with the trends we observed in the shapes of psychophysical tuning curves determined for several values of f_s in the range below 100 Hz (Jurado et al., 2010a).

Consider now the p values shown in Table II for the LF-emphasis and No-emphasis conditions (METF-separate analysis only). When no LF emphasis was applied, the value of p_l could not be reliably determined with f_s = 50 Hz, presumably because the lower band of noise made a

TABLE I. Parameters derived from fitting the mean masked thresholds for each value of f_S in the range 1000-80 Hz. TH: mean absolute threshold in dB SPL. The root-mean-square error (RMSe) is given in dB. Values of *r* and *K* are also expressed in dB. *A* is the asymmetry index. In all cases the spectrum level was 50 dB. Estimates of the ERB, obtained for the three cases described in the text, are shown in the three right-most columns. For f_S = 125 Hz, the mean of the parameters obtained from fitting the mean masking data of each group (Group 1 and Group 2) is given.

		METF-separate					System as a whole								
f_S	TH	p_u	p_l	*r*	*K*	RMSe	p_u	p_l	*r*	*K*	*A*	RMSe	ERB[a]	ERB[b]	ERB[c]
1000	0.8	29.9	22.5	-54.0	-5.2	1.2	30.4	23.2	-56.3	-5.0	1.3	1.2	156.0	152.1	156.0
500	5.2	27.5	20.6	-48.3	-6.4	1.2	27.1	22.2	-49.1	-6.0	1.2	1.1	84.9	82.1	84.9
250	10.2	20.0	15.2	-43.5	-5.2	1.0	20.0	17.9	-41.0	-4.2	1.1	0.9	58.1	53.0	58.1
125	18.0	18.3	10.7	-39.9	-0.2	0.5	16.9	14.7	-40.9	1.2	1.1	0.8	37.3	31.8	37.2
100	24.0	18.4	10.6	-37.2	0.2	0.7	16.6	14.6	-37.1	0.8	1.1	0.8	29.7	25.8	29.5
80	25.9	14.5	9.6	-31.8	-0.7	0.6	12.7	14.1	-31.2	-0.2	0.9	0.7	27.7	24.0	27.2

negligible contribution to the masking of the signal. The fitting procedure returned a value of p_l close to zero, but other values gave equally good fits. The data obtained with LF emphasis are probably more useful for determining the characteristics of both sides of the auditory filter. For these data, the estimated filter for f_s = 63 Hz was roughly symmetric, while for f_s = 50 Hz it had a steeper upper skirt. The *p* values obtained for the analysis of the combined data (with and without LF emphasis) were similar to those obtained for the LF-emphasis data alone, except that the value of p_l for f_s = 63 Hz was slightly greater for the combined data. The value of p_l for f_s = 50 Hz was approximately equal to 4. However, when the METF was assumed to have the form specified in ANSI (2007), the value of p_l was approximately 2. This makes sense, as, for very low frequencies, the ANSI-METF is steeper than the 'default' METF assumed here. When the METF is assumed to be steeper, then less steepness is assigned to the low-frequency skirt of the auditory filter. The sharpness of the low-frequency skirt of the auditory filter for low CFs clearly depends strongly on the assumed METF.

A further analysis of the estimated effect of the METF on tuning is given below.

2. Values of the ERB

Values of the ERB were calculated for three analysis cases: (a) METF separate; (b) System as a whole; (c) A composite case where the METF was included as part of frequency selectivity only for frequencies below 50 Hz, in the region where the shunt effect of the helicotrema is dominant. In this case, in the region below 50 Hz, the auditory filter shape was obtained by cascading the METF with the auditory filter derived using the METF-separate analysis. Above 50 Hz, the auditory filter shape derived using the METF-separate analysis was preserved. ERB values for cases (a) and (b) were calculated from the full expression for the filter shape (Eq. 2), taking the value of *r* into account. ERB values for (c) were obtained from numerical integration of the filter shape between g = –1 and g = 1. The results are shown in the right-most three columns of Table I and II.

For f_s from 1000 to 80 Hz, the ERB values decreased progressively with decreasing f_s and did not differ greatly for the three cases described above.

TABLE II. As Table I, but for f_S = 50 and 63 Hz. Results for conditions with and without LF emphasis are shown. Results for the combined data of the two conditions are also shown. For f_S = 50 Hz, the data were also analyzed using the ANSI-METF. The results of these analyses are shown in the rows labeled 'ANSI'.

			METF-separate					System as a whole								
	f_S	TH	p_u	p_l	r	K	RMSe	p_u	p_l	r	K	A	RMSe	ERB[a]	ERB[b]	ERB[c]
	63	34.8	10.3	4.9	-22.7	0.9	0.5	8.0	10.4	-22.5	1.1	0.8	0.6	37.9	28.4	30.2
No-emphasis	50	41.3	8.4	0	-13.0	-0.9	0.7	5.7	10.3	-9.6	-1.5	0.6	0.6	63.9	34.9	25.7
ANSI	50	--	8.6	0	-11.5	-0.5	0.7	--	--	--	--	--	--	64.3	--	25.5
	63	--	7.2	7.7	-26.4	-2.8	0.7	5.9	12.5	-26.2	-2.2	0.5	1.1	34.1	30.9	31.5
LF-emphasis	50	--	7.4	3.7	-10.3	-2.7	0.7	3.3	10.4	-22.0	-2.1	0.3	1.0	44.6	37.6	27.1
ANSI	50	--	6.6	1.9	-8.4	-3.0	0.6	--	--	--	--	--	--	59.5	--	30.8
	63	--	7.1	10.5	-35.2	-1.0	1.2	--	--	--	--	--	--	29.6	--	28.6
Combined	50	--	7.3	3.7	-10.9	-2.4	1.0	--	--	--	--	--	--	43.6	--	26.6
ANSI	50	--	7.2	1.7	-9.1	-2.8	1.0	--	--	--	--	--	--	59.0	--	29.2

ERBs obtained for case (c) were almost identical to those obtained for the METF-separate case, suggesting that the helicotrema has little or no effect in determining frequency selectivity at these values of f_s. For f_s = 63 and 50 Hz, the ERB values differed more across the three cases. This was especially so for the condition without LF emphasis, but, as noted earlier, for that condition the low-frequency side of the filter was not well defined. For the LF-emphasis condition, the ERB values were generally greater for the METF-separate analysis than for the System-as-a-whole analysis. This makes sense, as treating the METF as part of the auditory filter has the effect of increasing the steepness of the lower skirt and decreasing the estimated ERB. The largest differences in ERBs between the METF-separate case and case (c) occurred at 63 and 50 Hz. This suggests that the helicotrema increasingly influences frequency selectivity for CFs that approach the apical end of the cochlea.

It is noteworthy that, for cases (a) and (b) (i.e. METF-separate and System-as-a-whole analyses), the ERB increased as f_s decreased from 63 to 50 Hz. The increase in ERB is not predicted by Eq. 1. For case (a) the increase was mainly caused by a flattening of the lower skirt of the filter, probably related to the influence of the helicotrema on tuning, as described above. For case (b) a decrease in slope of the upper skirt of the filter also contributed to the increase in the ERB. This considerable flattening of the upper skirts, and related increase in ERB, has also been observed for psychophysical tuning curves for f_s values in the range below 100 Hz (Jurado *et al.*, 2010a).

The observed increase in bandwidth may reflect a genuine increase in auditory filter bandwidth at very low CFs. However, it may have occurred because the CF at the extreme apical end of the basilar membrane is a little above 50 Hz. If that were the case, then all stimuli for the conditions with f_S = 50 Hz would have been perceived via auditory filters with CFs above 50 Hz. To assess this possibility, the data for f_S = 50 Hz were re-analyzed, assuming that the lowest auditory filter was located a little above 50 Hz (e.g. 55 or 60 Hz) and not allowing it to shift below this limit. This led to larger RMSe values than when the CF was allowed to vary freely. This outcome suggests that it was reasonable to assume that there are auditory filters with CFs down to about 50 Hz.

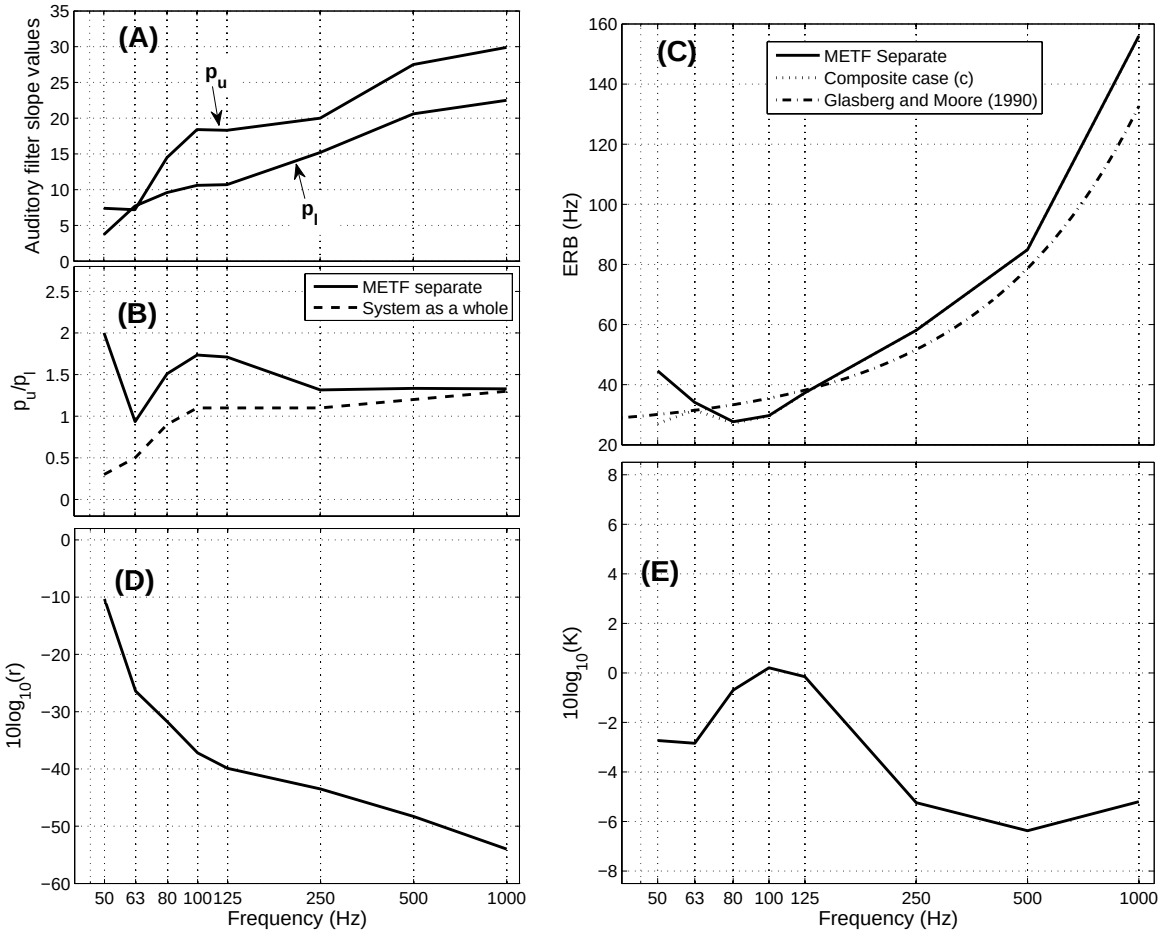

FIG. 5. (A): derived auditory filter slope parameters (p values) for the METF-separate analysis; (B): asymmetry parameter, *A*, for the METF-separate and System-as-a-whole analyses (solid and dashed lines, respectively); (C): ERB values obtained for the METF-separate analysis (solid line), the analysis treating the METF as part of frequency selectivity for frequencies below 50 Hz ('composite' case, dotted line). The ERB values predicted by Eq. 1 are plotted as the dashed-dotted line; (D): values of the dynamic range parameter, *r* (dB) for the METF-separate analysis; (E): values of the detection efficiency parameter, *K* (dB), for the METF-separate analysis. In all cases (A to E) for f_S = 63 and 50 Hz, values are based on the data obtained with LF emphasis.

However, the failure to find a decrease in RMSe when the filter was not allowed to shift below a certain limit might reflect inadequacies in the assumed filter shape used to fit the data. Also, our analyses assumed a smoothly varying METF, but there may in fact be irregularities in the METFs of individual subjects at very low CFs (Marquardt and Pedersen, 2010). Jurado *et al*. (2010b) showed that such irregularities, which are thought to be related to the action of the helicotrema (Marquardt and Hensel, 2008), may account for the broadening of the upper side of the auditory filter at very low CFs. Nevertheless, both the increase in ERB for very low values of f_S and the assumption that there are auditory filters with CFs down to 50 Hz are consistent with the psychophysical tuning curves measured by Jurado *et al*. (2010a) for values of f_S below 100 Hz.

Another possible reason for the increase of the ERB at very low CFs, but one still related to the action of the helicotrema, derives from the work of Schick (1994). He estimated the influence of the helicotrema in determining the location of the maximum amplitude of cochlear traveling waves. His results suggest that the helicotrema has to meet conflicting requirements. On the one hand, a very large helicotrema (leading to low impedance) will act as an ideal short circuit. This reduces sensitivity to intense low-frequency sounds, reducing masking effects and providing protection from large but slow changes in air pressure. On the other hand, a very large helicotrema would decrease the change in position of maximum displacement of the traveling wave produced by a given change of the frequency of the input, effectively decreasing frequency resolution, particularly at low frequencies. It is possible that the relatively large helicotrema in humans has evolved to favor a more efficient reduction in sensitivity to low-frequency sound at the expense of a worsening of frequency selectivity at very low frequencies.

Figure 5(C) compares ERB values predicted by Eq. 1 (Glasberg and Moore, 1990, dashed-dotted line) with the values obtained using the METF-separate analysis (solid line), and the analysis treating the METF below 50 Hz as part of the auditory filter (case (c), dotted line). The ERB values are generally similar to those predicted by Eq. 1, except for f_S = 1000 Hz, where the ERB values are slightly greater than predicted by Eq. 1, probably because the level/ERB_N of the noise used here increased with increasing f_S (as the spectrum level was fixed), and Eq. 1 is intended to apply for a level of 51 dB/ERB_N. For f_S = 50 Hz, the ERB value is above that predicted by Eq. 1 for the METF-separate analysis, but is slightly below that predicted by Eq. 1 for the analysis treating the METF below 50 Hz as part of the auditory filter.

3. Dynamic range, r

The values of the dynamic range parameter, *r*, derived using the METF-separate analysis are shown in Fig. 5(D). As expected, the dynamic range observed in the masking data was reflected in that of the derived auditory filters. The values of *r* decreased monotonically with decreasing f_S. The decrease in *r* with decreasing f_S corresponded well with the increase in absolute thresholds, and the correlation between *r* and the absolute threshold at f_S was significant (σ = 0.98, $p < 0.01$). Although the correlation is much higher than observed previously, the trend is consistent with the results of Patterson *et al*. (1982). The decrease in *r* was especially marked for f_S below 125 Hz. As noted earlier, the dynamic range of the data and of the derived filters was not directly limited by the absolute threshold at f_S, since the lowest masked thresholds were always well above the absolute threshold. However, the masked thresholds may have been affected by the approach to absolute threshold (Humes and Jesteadt, 1989).

In addition to the analysis in which *r* was constrained to be the same for the two sides of the filter, we tried fitting the data with a different value of *r* for each side of the filter. This did not improve the fits, consistent with the idea that the values of *r* were constrained by the approach to absolute threshold. The values of *r* for the System-as-a-whole analysis were generally similar to those obtained in the METF-separate analysis, except for the case with LF emphasis and f_S = 50 Hz, where *r* was markedly smaller for the METF-separate analysis.

C. Detection efficiency, *K*

Values of the detection efficiency parameter, *K*, are shown in Fig. 5(E) for the METF-separate analysis. For f_S between 100 and 500 Hz, the values of *K* increased (indicating poorer efficiency) with decreasing f_S. This trend is consistent with previous studies (Moore *et al.*, 1990; Unoki *et al.*, 2006a). However, for f_S below 100 Hz, the values of *K* tended to decrease, suggesting a moderate improvement in detection efficiency at very low CFs. A similar overall trend was observed for the System-as-a-whole analysis, with values slightly above those shown in Fig. 5(E). The improvement in detection efficiency for very low CFs may be a result of the ability to "listen in the dips" of the masker when the fluctuations in amplitude of the masker become very slow (Zwicker and Schutte, 1973; Moore and Glasberg, 1987a). It may also have been the case that, when the fluctuations were very slow, subjects could remember the pattern of amplitude fluctuation and pick the signal interval as the one in which the pattern of fluctuation was different from that in the other intervals; recall that the noise (and its pattern of amplitude fluctuation) was the same for all three intervals of a forced-choice task. However, as noted earlier, a pilot study showed that masked thresholds were similar when a "running" noise was used, so it seems unlikely that this effect was large.

D. Derived auditory filter shapes

Derived auditory filter shapes are shown in Fig. 6. The upper panel shows results from the METF-separate analysis (solid lines). Relative to the CF, the filter shapes tend to broaden as f_S decreases, especially in the range below 100 Hz, where the low-frequency skirts of the filters flatten off considerably. Although the latter effect is much more marked than observed previously for CFs above 100 Hz, it is consistent with previous observations by Baker and Rosen (2006). The broadening of the low-frequency sides of the filters appears to be a consequence of excluding the filtering effect of the METF from auditory tuning. The dotted line for f_S = 50 Hz shows the filter shape derived using the ANSI-METF. The lower skirt of this filter is even shallower than that of the filter derived using the 'default' METF. The thin-dashed lines show the low-frequency tails of the filters obtained when the effects of the METF below 50 Hz were treated as part of auditory filtering. As a natural consequence of the highpass nature of the METF, the low-frequency sides of these 'composite' filters sharpen considerably.

It is of interest to compare the filter shapes derived from the System-as-a-whole analysis with those obtained by cascading the assumed METF with the filter shapes obtained using the METF-separate analysis. These are shown in the lower panel of Fig. 6 for f_S = 50 and 63 Hz (solid lines: METF cascaded with derived auditory filter; dashed lines: System-as-a-whole fit). The filter shapes for these two cases were very similar except for frequencies well below the CF, for which the filters for the System-as-a-whole case had shallower slopes. This is a consequence of modeling the filter with a "hard" dynamic range parameter, *r*, which flattens the low-frequency skirt of the roex filter; the 'combined-effects' filter is less influenced by *r*, due to the much steeper slope of the METF.

The overall shapes of the filters and the degree to which the METF influences frequency selectivity are consistent with observations based on psychophysical tuning curves for low values of f_S (Jurado et al., 2010a).

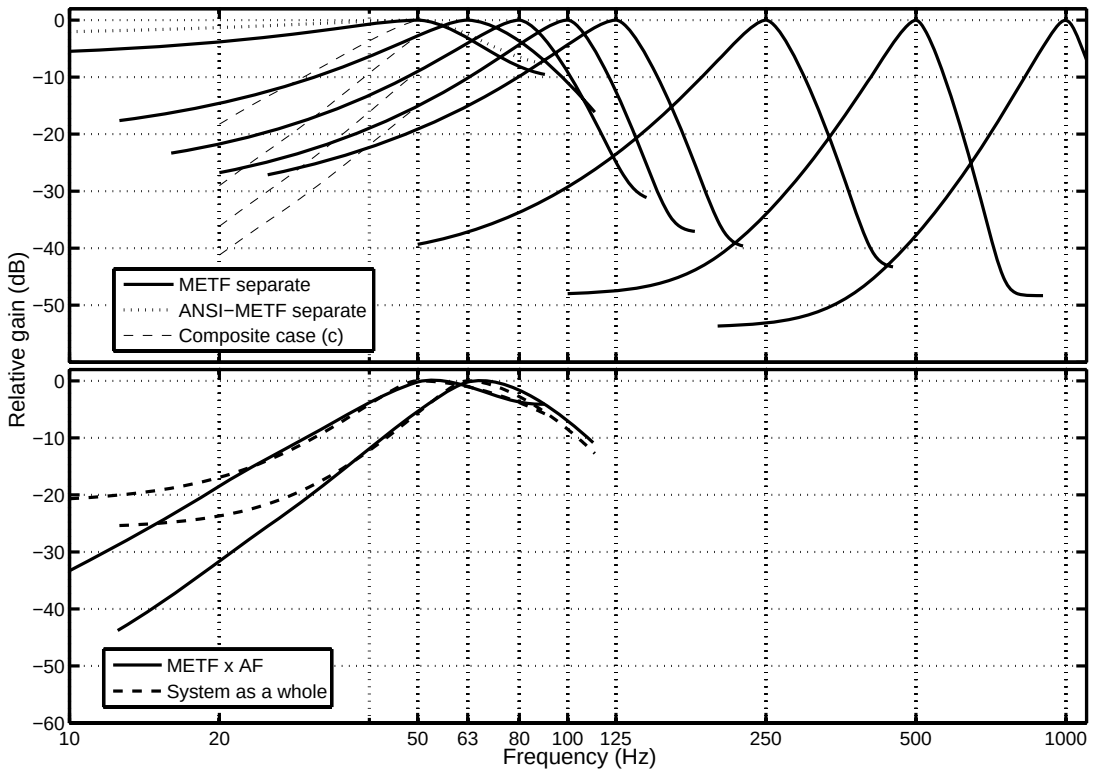

FIG. 6. Derived auditory filter shapes for f_S from 50 to 1000 Hz (LF-emphasis conditions only shown for f_S = 50 and 63 Hz). The upper panel shows the filter shapes corresponding to the METF-separate analysis (solid line) and the low-frequency tails of the filters after including the region below 50 Hz in the METF as part of auditory tuning (shown in thin-dashed line only for f_S below 100 Hz). The dotted line for f_S = 50 Hz is the filter shape derived using the ANSI-METF. The lower panel shows 'combined' filter shapes obtained for f_S = 50 and 63 Hz, after cascading the METF with the auditory shapes derived from the METF-separate analysis (METF x AF; solid line), and the roex filter shapes obtained in the System-as-a-whole fit (dashed line).

V. GENERAL DISCUSSION

A. Comparison with earlier results

Peters and Moore (1992) obtained estimates of auditory filter shape using the notched-noise method for CFs of 100, 200, 400 and 800 Hz. The overall noise level was 77 dB SPL. The data were analyzed in a way similar to our METF-separate analysis, but using an METF based on the shape of the 100-phon equal-loudness contour. For comparable CFs, their ERB values for normal-hearing subjects were generally close to but slightly larger than ours for the METF-separate analysis. Their values of p_u and p_l were also close to ours. Their value of K for f_S = 800 Hz was similar to ours for f_S = 1000 Hz, but their K values were generally about 4 dB larger than ours. Nevertheless, their values increased with decreasing CF, as ours did between 100 and 500 Hz.

Rosen and Stock (1992) obtained estimates of auditory filter shape using symmetric notched noise (five notch widths per f_S) for f_S values of 125, 250, 500 and 1000 Hz. They used N_0 values of 40, 50, 60 and 70 dB. The data were analysed in a way similar to our System-as-a-whole case. For

N_0 = 50 dB, the same as used here for the conditions without LF emphasis, their *p* values were somewhat smaller than ours and their ERB values (inferred from the 3-dB bandwidths reported by them) were somewhat larger than ours. However, their ERB values at higher CFs were also larger than typically reported in other studies (Glasberg and Moore, 1990). Their *K* values were very close to those found by us for f_S down to 125 Hz (less than 1 dB difference on average), and their values, like ours, increased with decreasing f_S from 125 Hz to 500 Hz.

Overall, we conclude that our data are reasonably consistent with those of comparable studies that included low signal frequencies (down to 100 Hz).

B. The effect of including the METF when deriving auditory filter shapes

Our results suggest that the METF increasingly influences frequency selectivity as f_S is decreased, particularly influencing the sharpness of the low-frequency skirt of the auditory filter. This effect was especially marked for f_S = 50 and 63 Hz; treating the METF as a fixed filter prior to auditory filtering resulted in an increase in ERB values as f_S was reduced from 63 to 50 Hz. Models of cochlear mechanics indicate that, for frequencies below about 50 Hz, the traveling wave extends along almost the whole length of the basilar membrane (BM), and pressure differences across the BM are partially shunted by the helicotrema (Schick, 1994; Marquardt and Hensel, 2008). This cochlear mechanism significantly influences tuning for CFs that approach the apical end of the cochlea. Indeed the effective tuning of the low-frequency side of the auditory filter for f_S = 50 Hz seems to be determined to a large extent by the shunt action of the helicotrema.

Results from our study measuring psychophysical tuning curves (Jurado *et al.*, 2010a) agree with this notion. For a 'composite' auditory filter that includes the region of the METF below 50 Hz as part of auditory filtering (case (c) described above), the ERB did not increase when f_S was decreased from 63 to 50 Hz, but remained roughly constant at a value close to that predicted by Eq. 1. The METF can, nevertheless, still be considered to be a fixed and level-independent filter, and, with the possible exception of the region below about 50 Hz, it makes sense to consider it as a filter applied prior to auditory filtering.

VI. CONCLUSIONS

We estimated the characteristics of auditory filters using the notched-noise method for values of f_S from 50 to 1000 Hz. The following are the main conclusions:

(1) Increasing the relative level of the lower masker band made it possible to define the low-frequency skirt of the auditory filter for very low values of f_S.

(2) Detection efficiency worsened with decreasing f_S for f_S between 100 and 500 Hz. For f_S below 100 Hz, a moderate improvement in detection efficiency was observed.

(3) Frequency selectivity for CFs approaching the apical end of the cochlea, such as 63 and 50 Hz, was strongly affected by the form of the (assumed) METF below 50 Hz, which is thought to reflect the shunting effect of the helicotrema.

(4) The sharpness of the skirts of the derived filters decreased progressively with decreasing f_S.

(5) For the analysis of the System as a whole, the asymmetry of tuning reversed with decreasing f_S; for f_S = 100 Hz and above, the upper skirts were steeper than

the lower skirts, while for lower values of f_S the reverse was true.

(6) The dynamic range of the derived filters decreased with decreasing f_S. This decrease was particularly sharp in the range below 125 Hz and it closely corresponded to the rapid increase in absolute thresholds in this range. The decrease may not be an inherent property of the auditory filters, but may reflect the approach of masked thresholds to absolute threshold for large notch widths.

(7) The ERBs of the derived filters decreased with decreasing f_S for f_S down to 80 Hz. For lower values of f_S, the ERBs depended strongly on assumptions made about the METF.

ACKNOWLEDGMENTS

We thank Christian S. Pedersen and Henrik Møller for helpful advice concerning the design of the experiments and Claus Vestergaard and Peter Dissing for their help with the experimental set-up. We also thank the Associate Editor Chris Plack and two anonymous reviewers for helpful comments on an earlier version of this paper.

REFERENCES

Aibara, R., Welsh, J. T., Puria, S., and Goode, R. L. (**2001**). "Human middle-ear sound transfer function and cochlear input impedance," Hear. Res. **152**, 100-109.

ANSI (**2007**). *ANSI S3.4-2007. Procedure for the computation of loudness of steady sounds* (American National Standards Institute, New York).

Baker, R. J., and Rosen, S. (**2006**). "Auditory filter nonlinearity across frequency using simultaneous notched-noise masking," J. Acoust. Soc. Am. **119**, 454-462.

Bos, C. E., and de Boer, E. (**1966**). "Masking and discrimination," J. Acoust. Soc. Am. **39**, 708-715.

Cheatham, M. A., and Dallos, P. (**2001**). "Inner hair cell response patterns: implications for low-frequency hearing," J. Acoust. Soc. Am. **110**, 2034-2044.

Dallos, P. (**1970**). "Low-frequency auditory characteristics: species dependence," J. Acoust. Soc. Am. **48**, 489-499.

Dallos, P. (**1973**). *The Auditory Periphery: Biophysics and Physiology* (Academic Press, New York), pp. 548.

de Boer, E., and Bouwmeester, J. (**1974**). "Critical bands and sensorineural hearing loss," Audiology **13**, 236-259.

Fletcher, H. (**1940**). "Auditory patterns," Rev. Mod. Phys. **12**, 47-65.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103-138.

Glasberg, B. R., and Moore, B. C. J. (**2000**). "Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise," J. Acoust. Soc. Am. **108**, 2318-2328.

Glasberg, B. R., Moore, B. C. J., Patterson, R. D., and Nimmo-Smith, I. (**1984**). "Dynamic range and asymmetry of the auditory filter," J. Acoust. Soc. Am. **76**, 419-427.

Hensel, J., Scholz, G., Hurttig, U., Mrowinski, D., and Janssen, T. (**2007**). "Impact of infrasound on the human cochlea," Hear. Res. **233**, 67-76.

Humes, L. E., and Jesteadt, W. (**1989**). "Models of the additivity of masking," J. Acoust. Soc. Am. **85**, 1285-1294.

ISO 226 (**2003**). *Acoustics - normal equal-loudness contours* (International Organization for Standardization, Geneva).

Jackson, L. B. (**1996**). *Digital Filters and Signal Processing* (Kluwer Academic Publishers, Boston), pp. 520.

Jurado, C. A., Moore, B. C. J., and Pedersen, C. S. (**2010a**). "Psychophysical tuning curves for very low centre frequencies," in *Proceedings of 20th International Congress on Acoustics, ICA 2010* (Sydney, Australia), CD-ROM.

Jurado, C. A., Pedersen, C. S., and Marquardt, T. (**2010b**). "Frequency selectivity at very low

centre frequencies: the influence of the helicotrema on individual differences in low-frequency sound perception," in *Proceedings of 14th International Conference on Low Frequency Noise and Vibration and its Control* edited by F. Christensen, and C. S. Pedersen (Aalborg, Denmark), pp. 129-145.

Leventhall, H. G. (**2004**). "Low frequency noise and annoyance," Noise and Health **6**, 59-72.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467-477.

Marquardt, D. (**1963**). "An algorithm for least-squares estimation of nonlinear parameters," SIAM J. Appl. Math **11**, 431-441.

Marquardt, T., and Hensel, J. (**2008**). "A lumped-element model of the apical cochlea at low frequencies," in *Concepts and Challenges into the Biophysics of Hearing*, edited by N. P. Cooper, and D. T. Kemp (World Scientific Publishing Co, London), pp. 337-339.

Marquardt, T., and Pedersen, C. S. (**2010**). "The infuence of the helicotrema on low-frequency hearing," in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York), pp. 25-36.

Marquardt, T., Hensel, J., Mrowinski, D., and Scholz, G. (**2007**). "Low-frequency characteristics of human and guinea pig cochleae," J. Acoust. Soc. Am. **121**, 3628-3638.

Møller, H., Hammershøi, D., Jensen, C. B., and Sørensen, M. F. (**1995**). "Transfer characteristics of headphones measured on human ears," J. Audio Eng. Soc. **43**, 203-217.

Moore, B. C. J., and Glasberg, B. R. (**1987a**). "Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes," J. Acoust. Soc. Am. **82**, 69-79.

Moore, B. C. J., and Glasberg, B. R. (**1987b**). "Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns," Hear. Res. **28**, 209-225.

Moore, B. C. J., and Sek, A. (**1995**). "Auditory filtering and the critical bandwidth at low frequencies," in *Advances in Hearing Research*, edited by G. A. Manley, G. M. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 425-436.'

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). "A model for the prediction of thresholds, loudness and partial loudness," J. Audio Eng. Soc. **45**, 224-240.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1990**). "Auditory filter shapes at low center frequencies," J. Acoust. Soc. Am. **88**, 132-140.

Patterson, R. D. (**1976**). "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. **59**, 640-654.

Patterson, R. D., and Moore, B. C. J. (**1986**). "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 123-177.

Patterson, R. D., and Nimmo-Smith, I. (**1980**). "Off-frequency listening and auditory filter asymmetry," J. Acoust. Soc. Am. **67**, 229-245.

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). "The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold," J. Acoust. Soc. Am. **72**, 1788-1803.

Pedersen, C. S., and Marquardt, T. (**2009**). "Individual differences in low-frequency noise perception," in *Proceedings of INTER NOISE 2009* (Ottawa, Canada).

Peters, R. W., and Moore, B. C. J. (**1992**). "Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects," J. Acoust. Soc. Am. **91**, 256-266.

Puria, S., Rosowski, J. J., and Peake, W. T. (**1997**). "Sound-pressure measurements in the cochlear vestibule of human-cadaver ears," J. Acoust. Soc. Am. **101**, 2754-2770.

Rife, D. D., and Vanderkooy, J. (**1989**). "Transfer-function measurement with maximum-length sequences," J. Audio Eng. Soc. **37**, 419-444.

Rosen, S., and Stock, D. (**1992**). "Auditory filter bandwidths as a function of level at low frequencies (125 Hz-1 kHz)," J. Acoust. Soc. Am. **92**, 773-781.

Rosen, S., Baker, R. J., and Darling, A. (**1998**). "Auditory filter nonlinearity at 2 kHz in

normal hearing listeners," J. Acoust. Soc. Am. **103**, 2539-2550.

Rosowski, J. J. (**1996**). "Models of external- and middle-ear function," in *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (Springer, New York), pp. 15-61.

Schick, F. (**1994**). "The helicotrema and the frequency resolution in the inner ear," Acustica - Acta Acustica **80**, 463-470.

Shaw, E. A. G. (**1974**). "Transformation of sound pressure level from the free field to the eardrum in the horizontal plane," J. Acoust. Soc. Am. **56**, 1848-1861.

Shera, C. A., Guinan, J. J., Jr., and Oxenham, A. J. (**2002**). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proc. Natl. Acad. Sci. USA **99**, 3318-3323.

Unoki, M., Kazuhito, I., Ishimoto, Y., and Tan, C. T. (**2006a**). "Estimate of auditory filter shape using notched-noise masking for various signal frequencies," Acoust. Sci. Tech. **27**, 1-11.

Unoki, M., Irino, T., Glasberg, B. R., Moore, B. C. J., and Patterson, R. D. (**2006b**). "Comparison of the roex and gammachirp filters as representations of the auditory filter," J. Acoust. Soc. Am. **120**, 1474-1492.

Zwicker, E., and Schutte, H. (**1973**). "On the time pattern of the threshold of tone impulses masked by narrow band noise," Acustica **29**, 343-347.

Psychophysical tuning curves for frequencies below 100 Hz

Carlos Jurado[a)] and Christian S. Pedersen
Section of Acoustics, Department of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7-B5, 9220 Aalborg Ø, Denmark

Brian C.J. Moore
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

[a)] email: cjo@es.aau.dk

Abstract

Psychophysical tuning curves (PTCs) were measured for sinusoidal signals with frequency f_s = 31.5, 40, 50, 63, and 80 Hz, using sinusoidal and narrowband-noise maskers. For the former, conditions were included where a pair of beating tones were added to reduce the use of cues related to beats. Estimates of each subject's middle-ear transfer function (METF) were obtained from equal-loudness contours measured from 20 to 160 Hz. With decreasing f_s, the PTCs progressively broadened and became markedly asymmetrical, with shallow upper skirts and steep lower skirts. For the sinusoidal maskers, the tips were more irregular than for narrowband noise maskers or when beating tones were added. For f_s = 31.5 and 40 Hz, the tips of the PTCs always fell above f_s. Allowing for the METF so as to infer underlying filter shapes resulted in flatter lower skirts, especially below 40 Hz, and reduced the frequency at the tips for f_s between 31.5 and 50 Hz; however, the tips did not fall below 40 to 50 Hz. The bandwidths of the PTCs increased with decreasing f_s below 80 Hz. However, bandwidths remained roughly constant if the METF was included as part of auditory filtering for frequencies below 40 Hz.

I. INTRODUCTION

Noise with energy in the low-frequency range (below about 200 Hz) is known to produce problems with annoyance and represents an environmental problem (Leventhall, 2004). While the audibility and loudness of sinusoidal signals can be predicted from known threshold curves and equal-loudness contours (ISO 226, 2003), environmental noises are often complex signals. To understand how these signals excite the hearing organ, and to model and understand human sound perception of low frequencies, the characteristics of auditory frequency selectivity are required (Moore *et al.*, 1997). However, few results exist that describe the characteristics of frequency selectivity at very low frequencies.

Jurado and Moore(2010) obtained estimates of the bandwidth and shape of the auditory filter for several center frequencies (CF) down to 50 Hz, using the notched-noise method (Patterson, 1976; Patterson and Nimmo-Smith, 1980; Glasberg and Moore, 1990). Their results indicated that, while the bandwidth of the auditory filter at low CFs is wider relative to CF than at higher CFs, the equivalent rectangular bandwidth (ERB) decreases with decreasing CF at least down to 80 Hz. However, the ERB increased with a further decrease in CF

to 50 Hz. This worsening of tuning occurred mainly because the derived filter shapes had very shallow upper skirts. For CFs below 100 Hz, the lower filter skirts were sharper than the upper skirts, a clear reversal of the asymmetry often observed at higher CFs (Glasberg *et al.*, 1984). The asymmetry at low CFs was attributed to the attenuation produced by the middle ear and reduction in sensitivity produced by the helicotrema shunt mechanism, which act as a highpass filter at low frequencies (as observed in the physiological estimates of Cheatham and Dallos, 2001).

The middle-ear transfer function (METF) is assumed to become steeper in slope below about 50 Hz, because it includes the shunt effect of the helicotrema (Marquardt et al., 2007; Marquardt and Hensel, 2008). Jurado and Moore (2010) showed that this slope had a very strong influence on estimates of the lower skirt of the auditory filter as the CF approached 50 Hz. If the METF was treated as a fixed filter occurring before auditory filtering, the lower skirt was very shallow. Jurado et al. (2010) calculated auditory filter shapes by assuming that the auditory filters were preceded by METFs as estimated by Marquardt and Pedersen (2010) from a distortion-product isomodulation technique.

These METFs have a small resonance feature that forms a transition region separating two distinct slope regions (see Marquardt and Pedersen, 2010). The results suggested that the irregular resonance feature in the METFs could flatten off the upper skirts of the auditory filters for very low CFs and produce a shift in their tips that was typically upwards in frequency.

In the present work, psychophysical tuning curves (PTCs) were obtained for signal frequencies (f_s) below 100 Hz, to obtain a direct and detailed description of frequency selectivity at very low frequencies. The values of f_s used were 31.5, 40, 50, 63, and 80 Hz. Both sinusoidal and narrow-band noise maskers were used to mask sinusoidal signals at those values of f_s.

When sinusoidal maskers are used, beats between the signal and masker can provide a powerful detection cue (Wegel and Lane, 1924; Egan and Hake, 1950; Greenwood, 1971; Moore et al., 1998; Alcántara et al., 2000; Kluk and Moore, 2004). To reduce the influence of beat detection on the results, conditions were included where pairs of beating tones were added to the main masker, to reduce the influence of beats via the phenomenon of modulation detection interference (MDI) (Yost and Sheft, 1989; Bacon and Moore, 1993). This technique has been used previously by Alcántara et al. (2000) and by Kluk and Moore (2004).

The influence of the METF on tuning was also studied, by estimating its shape from equal-loudness contours (ELCs) measured using closely spaced frequencies in the range below 160 Hz. Additional aims were to quantify the extent of individual differences and to identify the CF of the "bottom" auditory filter, the one with the lowest CF, assuming that the "end" of the cochlea is reached for a frequency which may be above the lowest frequency that can be heard (Moore et al., 1997). The results were expected to contribute to general understanding of frequency selectivity at very low frequencies and therefore to help in characterizing human perception of low-frequency noise.

II. METHOD

A. Threshold procedure

A PTC is the level of a narrowband masker required just to mask a signal that is fixed in frequency and level, plotted as a function of masker frequency, f_m. The signal was a sinusoid with frequency f_s presented at a fixed level of 15 dB sensation level (SL).

The masker level was adjusted using a 3-alternative forced-choice task with a 3-up 1-down adaptive procedure to estimate the 79% point on the psychometric function (Levitt, 1971). The silent interval between the three stimuli in a trial was 400 ms. The masker was presented in all three intervals, and the signal was presented in one randomly-chosen interval. Intervals in which the signal might occur were indicated by lights on the response box.

The task was to press a button representing the interval that contained the signal. Feedback in the form of a light indicating the correct answer was provided after each response. The procedure started with a simple 1-up 1-down rule for the first four presentations in order to rapidly approach the region of masked threshold. The masker started at a relatively low level, so that the signal was clearly audible. This level was selected after pilot tests. The step size started at 8 dB, was decreased to 4 dB after two reversals, and was decreased to 2 dB after two further reversals, where it remained.

A total of 12 reversals was obtained and the masked threshold was estimated as the average of the masker levels at the last eight. Two threshold estimates were obtained for each condition. If the difference between them was more than 3 dB, a third estimate was obtained. All estimates were averaged. Absolute thresholds were obtained in a similar manner (but with a 3-down 1-up procedure).

B. Stimuli for measurement of PTCs

1. PTCs derived with sinusoidal maskers

The masker was a sinusoid. The values of the relative separation, $\Delta = (f_m - f_s / f_s)$, between f_m and f_s depended on f_s. To define the tip, finer spacing was used around f_s. A total of nine values of Δ were used for each f_s. These were: −0.6, −0.4, −0.2, 0, 0.2, 0.4, 1.2, 2.2, and 3 for 31.5 Hz; −0.6, −0.5, −0.4, −0.2, 0, 0.2, 0.4, 1.2, and 2.2 for 40 Hz; −0.6, −0.5, −0.4, −0.2, 0, 0.2, 0.4, 0.8, 1.6 for 50 and 63 Hz; and −0.6, −0.5, −0.4, −0.2, 0, 0.2, 0.4, 0.8, and 1.2 for 80 Hz. Signals were 1.2-s long, including 0.2-s linear ramps at start and end. The masker had the same length and ramp characteristics as the signal. The masker and signal were gated synchronously. When the signal and masker had the same frequency, they were added in phase. However, in the results presented below, the masker level at threshold was adjusted to allow for the in-phase addition of signal and masker; the masker level was expressed as the level that would have been required if the power of the signal and masker was added (i.e., as if they had been added with a relative phase of 90°).

2. PTCs derived with sinusoidal maskers and MDI tones

To study the influence of beat detection on the PTCs, "MDI tones" were added to the masker in all intervals of a forced-choice trial. These tones consisted of a fixed- and a variable-frequency tone; the frequency of the fixed tone was $2.2f_s$, except for f_s = 80 Hz, where it was $2f_s$. The variable-frequency tone had the same value of Δ relative to the fixed tone as for the masker and signal. For example, if the masker had a frequency of 30 (or 70) Hz, and the signal had a frequency of 50 Hz, the fixed MDI tone had a frequency of 110 Hz and the variable MDI tone had a frequency of 90 (or 130) Hz. The frequencies of the MDI tones were chosen so that they would have little or no direct (energetic) masking effect on the signal.

For the first two subjects tested (subjects 7 and 8), the MDI tones were included for Δ = −0.4, −0.2, 0.2, and 0.4. However, this appeared not to cover the whole range of

masker frequencies over which beat detection might influence the PTC, particularly for the lowest signal frequencies. Therefore, for the rest of the subjects, a value of Δ of 1.2, was included for f_s = 31.5 and 40 Hz, and a value of Δ of 0.8 was included for f_s = 50 and 63 Hz. The level of the MDI tones was set so that they were salient enough to provide effective interference with beat detection, while avoiding levels that might be high enough to directly mask the signal. The appropriate level was found after initial tests, for each subject at each f_s. The level varied from 7 to 30 dB above the ISO 226 (2003) absolute threshold levels for the frequencies of the MDI tones. The sensation level, relative to the ISO 226 thresholds, was the same for the two tones within each MDI pair. The MDI tones were gated synchronously with the masker, the time configuration of all signals being the same as for the sinusoidal maskers alone.

3. PTCs derived with narrowband-noise maskers

Narrowband noise with a bandwidth of $0.5f_s$ was used as a masker. This bandwidth was chosen to be wide enough to reduce the influence of beats as a cue (Moore et al., 1998; Alcántara et al., 2000; Kluk and Moore, 2004), while being narrow enough that it had only a small influence on the measured frequency selectivity. The values of Δ used were based on the center frequency of the masker and were the same as described earlier, with the exception that the largest value of Δ was 1.2 instead of 1.6 for f_s = 63 Hz.

The noise maskers were created by filtering a wideband white noise signal with a 200th-order Chebyshev type II digital infinite-impulse-response bandpass filter. The slopes of the filter were effectively infinite. Noise bursts were 1.3-s long, including 0.2-s linear ramps at the start and end. They were taken randomly from a 39-s noise buffer. In each 3-alternative sequence, the same randomly chosen burst was used. This was done to eliminate random variations in overall masker level from one interval to the next within a trial. The signal was 1.2-s long with 0.2-s ramps at start and end. The signal started 50 ms after the noise started, and ended 50 ms before the noise ended.

C. Measurement of equal-loudness contours

1. Stimuli

An ELC was obtained for each subject for frequencies between 20 and 160 Hz. This was done to obtain a rough estimate of the shape of the subject's METF. The reference signal had a frequency of 50 Hz and a fixed level of 40 dB SL. Comparison tones were used with frequencies of 20, 30, 35, 40, 45, 55, 60, 65, 70, 75, 80, 90, 100, 125, and 160 Hz. Their duration and ramp characteristics were the same as for the signals used with the sinusoidal masker.

2. Procedure

The reference (fixed level) and comparison (variable level) tones were presented in random order on each trial. The task was to indicate which of the two tones was the loudest, by pressing the button representing that tone. A 2-alternative task with a 1-up, 1-down adaptive loudness-balance procedure was used to estimate the 50% point on the psychometric function, corresponding to equal loudness of the two tones. Two interleaved tracks were used. The procedure randomly selected one of the two tracks when choosing the level of the next presentation. For a given comparison tone, the level of one track started 10 dB

below the 40 phon standardized ISO equal loudness level (ISO 226, 2003), and the other track started 10 dB above it (the ISO curve was interpolated to obtain values at all frequencies). This was done except when the frequency of the comparison tone was 20 Hz, in which case the tracks started at ±5 dB from the 40-phon level, avoiding initial levels that were too high. These starting levels were found to be adequate after pilot testing; typically the listener's point of subjective equality lay between these extremes.

The step size started at 8 dB, was decreased to 4 dB after 2 reversals, and was decreased to 2 dB after 2 further reversals, where it remained. For each track, 10 reversals were obtained and the point of subjective equality (PSE) was obtained from the average level at the last 6. For a given run (i.e. single measurement), one PSE was determined, corresponding to the average PSE of the two tracks. Two runs were performed, and the PSE was estimated from the average PSE for the two runs. If the two PSE estimates differed by more than 3 dB, a third run was performed and all three PSE estimates were averaged.

D. Subjects

A group of eight subjects, three female and five male were tested using the sinusoidal masker with and without MDI tones. Another group of eight subjects, four female and four male, were tested using the narrowband noise masker. Two subjects were common to the two groups for all values of f_s, while one subject who was tested with the sinusoidal masker was also tested with the noise masker but only at one signal frequency. ELCs were measured for all 14 subjects. Audiometric testing showed that all subjects had thresholds within the normal range (20 dB HL or better) for frequencies between 125 and 4000 Hz. Their ages were between 22 and 37 years.

E. Apparatus

A cabin especially designed for playback of low-frequency signals under pressure field conditions was used. The cabin had inner dimensions 0.8 × 1.4 × 0.9 m (~1 m3). It had four Seas 33 F-WKA 13-inch loudspeakers in each side wall, positioned behind covering panels. The listener sat facing the door, with the ears at 90° relative to the side walls. The cabin provided an effective pressure field within the overall volume up to about 61 Hz. However, deviations were small within a defined listening space up to higher frequencies (below ±3 dB up to about 150 Hz). A listener's position was defined by five points covering a region where it was expected that the listener's ears and center of the head would be located.

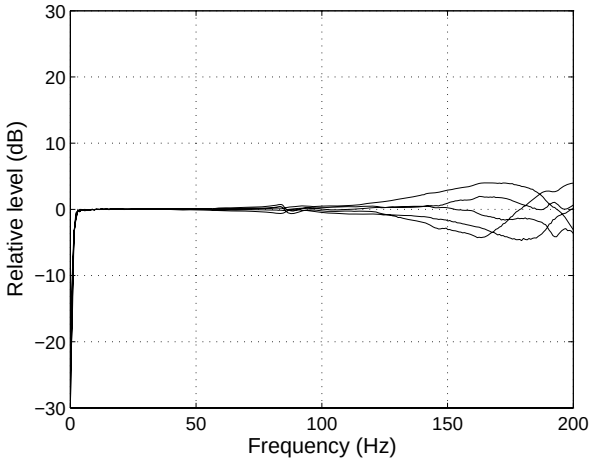

Figure 1. Differences between the frequency response at each of the five listening positions and the average response at these positions in the cabin.

Equalization was done by filtering all stimuli with a digital filter representing the inverse room magnitude response, obtained after averaging the magnitude responses measured at the five positions. When performing the task, all listeners were

instructed to find a comfortable position, without leaning to the sides, and to avoid moving much within blocks of measurements. Figure 1 shows the differences between the assumed frequency response (0 dB reference), and the frequency response at the five positions.

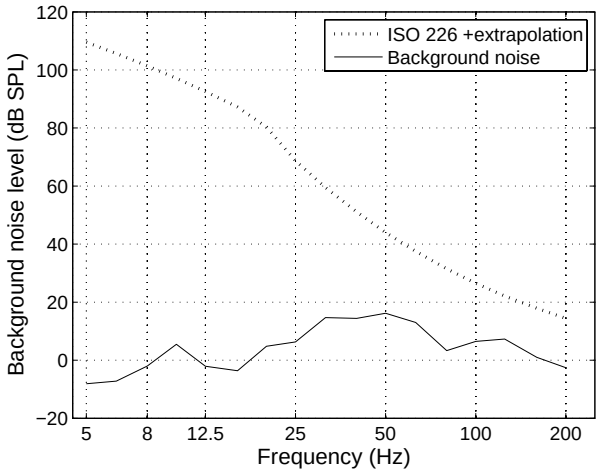

Figure 2. One-third octave background noise levels in the listening cabin (solid lines). The ISO 226-2003 (2003) absolute threshold levels are plotted as a dotted line. Below 20 Hz values correspond to those proposed by Møller and Pedersen(2004).

One-third octave background noise sound pressure levels were measured in the cabin. Levels were obtained over 250-s intervals and for worst-case scenarios, when activity was expected nearby in the building. The levels were measured with a low-noise microphone (Bruel & Kjær 4179, noise floor −5.5 dBA) using a 01 dB Harmonie system frequency analyzer. A worst-case scenario example is shown in Fig. 2. As illustrated, all third-octave noise levels in the low-frequency range were well below the hearing threshold. The overall background noise level in the cabin was 21 dB SPL (~0 dBA) over the frequency range 5-20,000 Hz.

Harmonic distortion was measured at 1/3-octave intervals using sinusoids with frequencies between 5 and 100 Hz, and at a level of 130 dB SPL. This level is much higher than the levels actually used in the experiment. In all cases, the 2nd, 3rd and 4th harmonics had levels at least 30, 40, and 50 dB, respectively, below that of the fundamental (typically by 10 dB more than this). Since the levels used in the experiment were typically at least 30 dB below 130 dB SPL, harmonic distortion levels were much lower than the levels described above. Thus, it seems likely that harmonic distortion in the acoustic signals was not audible.

Signals were sent through the ADAT optical outputs of an RME DIGI 96 PC soundcard to an RME ADI-8 D/A converter, using a sample rate of 48 kHz and 24-bit resolution. The eight loudspeakers were driven by a Crown Studio Reference I (1160 W) power amplifier.

III. RESULTS AND DISCUSSION

A. Parameters used for description of the PTCs

The location of the tip, the bandwidth, and asymmetry of the PTCs were quantified and used as the main descriptors. Additionally, the sharpness of the skirts was quantified in dB/octave. Slopes were measured between a point 5 dB above the tip to the value for the most extreme frequency used for that side of the PTC. The tip frequency was not taken as f_s, but rather was taken as the value of f_m for which the masker level at threshold was lowest. The bandwidths of the PTCs were determined by calculating both the equivalent rectangular bandwidth (ERB) and the 10-dB bandwidth of the PTCs. The ERB was obtained from numerical integration of linearly interpolated high-resolution versions of the PTCs. To calculate the ERB, each PTC was inverted and normalized to have 0-dB gain at its tip. The 10-dB bandwidth was also obtained from linearly interpolated versions of the PTCs, and was calculated using as a

reference point both the frequency of the tip of the PTC and the value of f_s.

To quantify the degree and direction of asymmetry of the PTCs, an asymmetry index, A_{10} dB, was defined as the ratio of the lower-side to upper-side 10-dB bandwidths, and was calculated relative to both the frequency at the tip and f_s. The index describes the asymmetry of the PTC in linear frequency units; its value is above 1 when the upper skirt is steeper and below 1 when the lower skirt is steeper. It is generally roughly consistent, both qualitatively and quantitatively, with the slope-ratio asymmetry index that has been used previously to describe auditory filter shapes (Glasberg et al., 1984; Jurado and Moore, 2010).

B. Sinusoidal masker without MDI tones

Figures 3 and 4 show the individual PTCs obtained using the sinusoidal masker for f_s = 31.5, 40, 50, 63, and 80 Hz (circles and solid lines). The PTCs are broad, reflecting the relatively poor frequency selectivity at these low frequencies. For f_s = 80 Hz, the tips of the PTCs were typically at f_s (in seven out of eight cases, see vertical dotted lines in Figs. 3 and 4). This was not the case for f_s = 31.5 and 40 Hz, where the tips of the PTCs were always above f_s. For f_s = 50 and 63 Hz, the tips were at f_s in one out of eight and three out of six cases, respectively, and were above it in all other cases (except for subject 3 at 50 Hz, for whom there was a distinct dip in the PTC at 25 Hz). The shifted tips for the lowest values of f_s may indicate that there are no auditory filters with CFs below about 50 Hz, as suggested by Moore et al. (1997). However, the steep slope of the METF may also lead to upward-shifted tips of the PTCs in this region (Jurado et al., 2010); this possibility is considered in section IV. For all values of f_s, distinct irregularities around f_s were observed in several cases, suggesting an influence of beat detection. For some subjects, such as 3, 5, and 6, these irregularities were evident for most values of f_s, while for others, such as 1 and 4, the PTCs were more regular in shape, particularly for the lower values of f_s. This may indicate that the salience of beats as a detection cue varied across subjects.

The masker levels at threshold typically increased with decreasing f_s, possibly as a consequence of the general increase in absolute thresholds with decreasing frequency. Mean absolute thresholds were 51.2, 43.5, 39.0, 33.1, and 26.3 dB SPL for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively. At f_s, the average difference between the signal level and the level of the sinusoidal masker at threshold was -11.1, -9.1, -9.7, -11.5, and -10.3 dB for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively; the masker level at threshold was always above the signal level. These differences are broadly consistent with those found for higher signal frequencies when sinusoidal maskers are used (Vogten, 1974; 1978).

For the mean PTCs obtained without MDI tones, the steepness in dB/oct of the lower and upper skirts (lower, upper) was about (19, 30), (22, 46), (18, 30), (15, 32), and (15, 23) for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively. The ratio of steepness for the upper and lower skirts was, on average, close to 1.4. However, on a linear frequency scale, the PTCs became increasingly asymmetric as f_s decreased, with steeper lower skirts than upper skirts. There were clear individual differences. For example, for subject 2, the upper skirt was very shallow for f_s from 31.5 to 50 Hz, while subject 1 had relatively steep upper skirts.

C. Sinusoidal masker with MDI tones

The results obtained in the presence of the MDI tones are shown by asterisks and

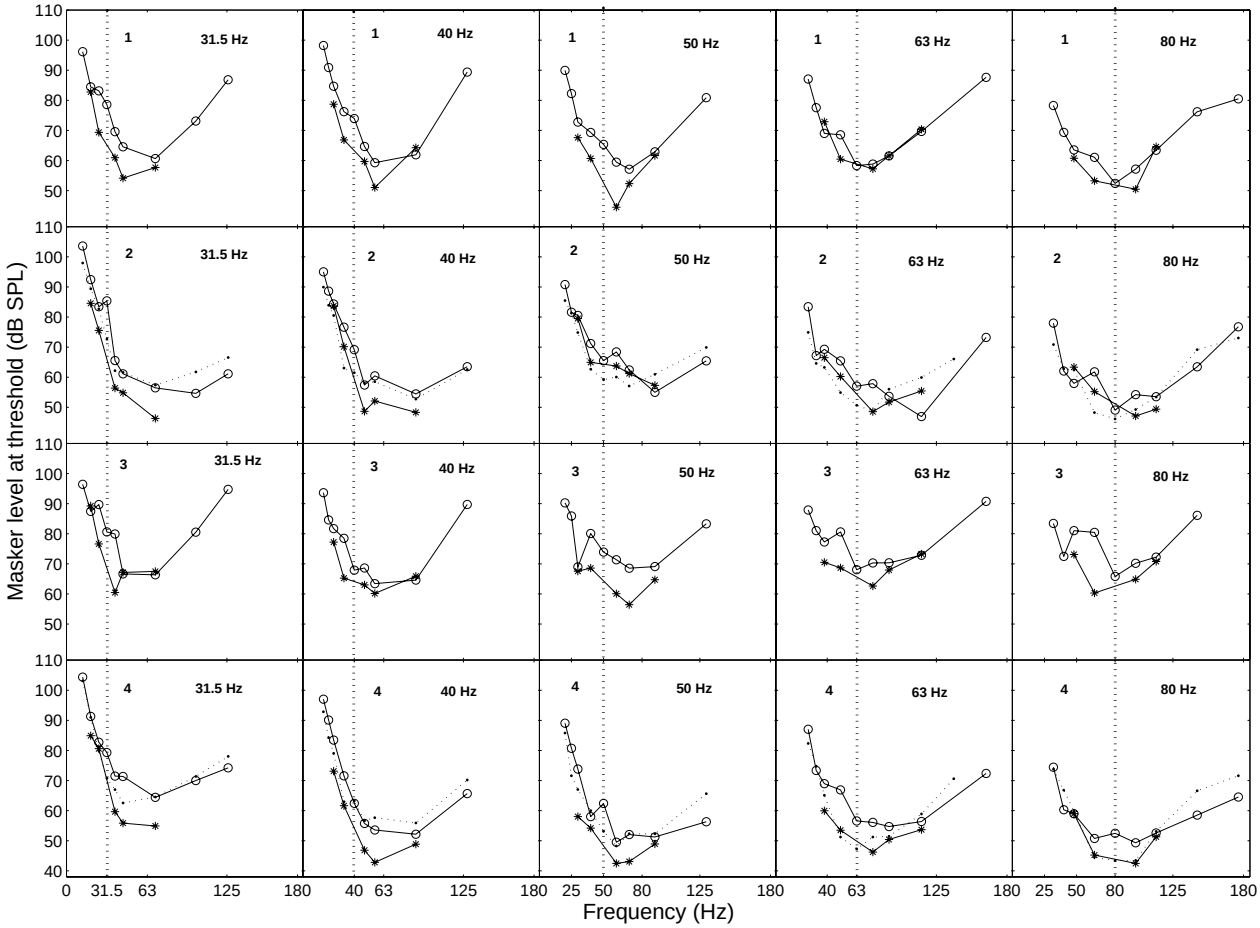

Figure 3. PTCs for f_s between 31.5 and 80 Hz, obtained using a sinusoidal masker without (circles and solid lines) and with (asterisks and solid lines) MDI tones for four subjects. Numbers in the upper left corner are the subject identifiers. When PTCs were obtained with the noise masker for the same subjects, they are plotted as dots and dotted lines. Vertical dotted lines indicate the value of f_s. No threshold was obtained for subject 3 at f_s = 80 and Δ =1.2 since the upper level limit of the system was reached for that frequency.

solid lines in Figs. 3 and 4. Masker levels at threshold were typically lower when the MDI tones were present than when they were absent. This suggests that beats played a significant role in detection of the signal for masker frequencies close to the signal frequency. However, lower masker levels might also have occurred if the MDI tones produced a large enough output from the auditory filter centered near f_s. The largest differences between the masker levels with and without MDI tones occurred for Δ = –0.2 and 0.2, being 11.3 dB and 9.4 dB, respectively, averaged across f_s. This is consistent with what has been observed at higher frequencies for similar values of Δ (Kluk and Moore, 2004; Alcántara et al., 2000).

For larger values of Δ, the effect produced by the MDI tones typically decreased. The average effect across f_s was: 4.4, and 5.8 dB for Δ = –0.4 and 0.4, respectively; 0.5 dB for Δ = 0.8 (f_s = 50 and 63 Hz); and 1.8 dB for Δ = 1.2 (f_s = 31.5 and 40 Hz). These results suggest that beat detection significantly influenced the masker level at threshold for Δ up to 0.4, while above Δ = 0.8 the effect of beat

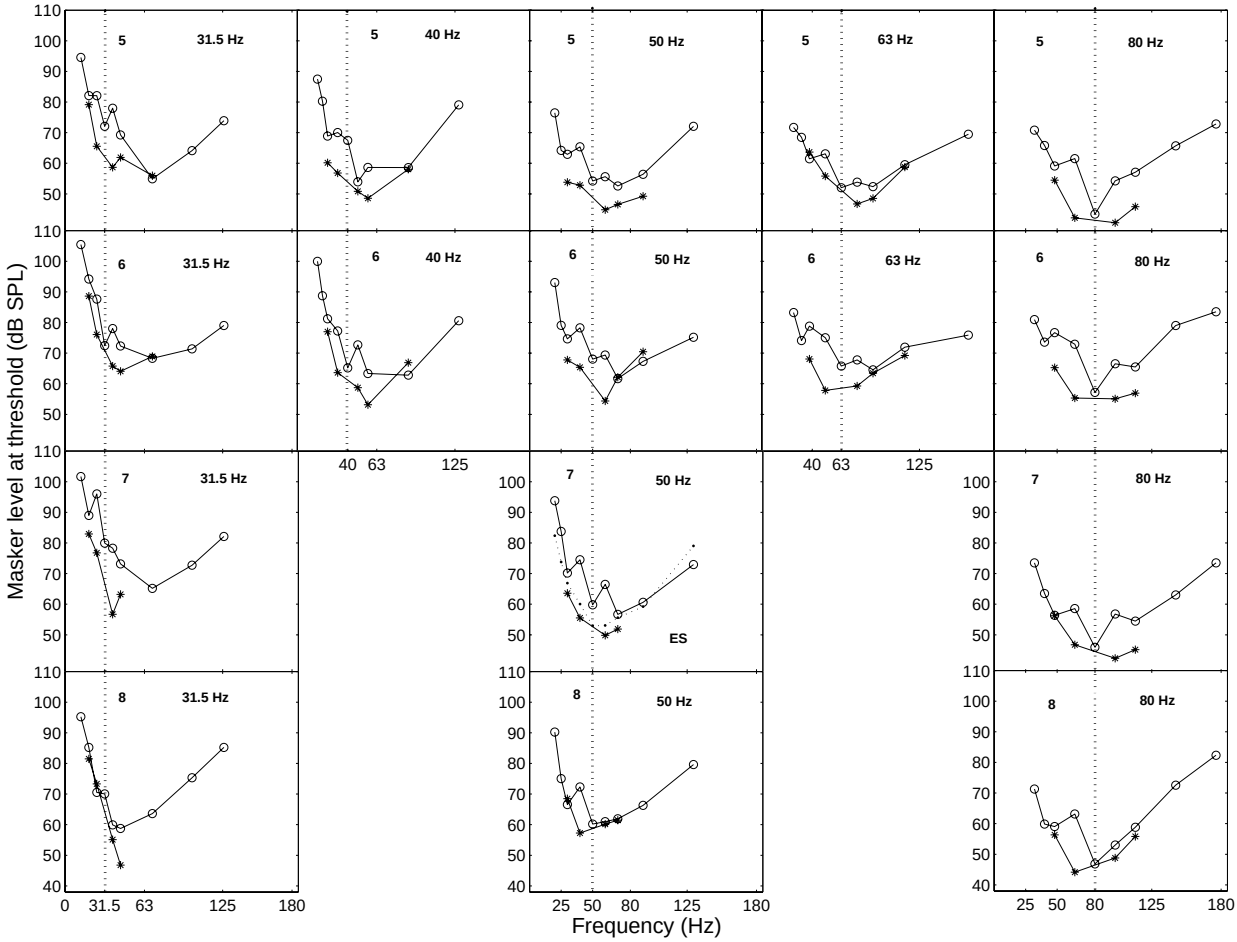

Figure 4. As Fig. 3, but for four more subjects. Two of the subjects were not tested for f_s = 40 and 63 Hz.

detection was marginal. For f_s = 80 Hz, the tips of the PTCs were generally more regular and broader when the MDI tones were present than when they were absent. This was also sometimes the case for f_s = 50 and 63 Hz (e.g., subjects 4 and 7 at 50 Hz; 1 and 5 at 63 Hz). However, sharp tips in the presence of the MDI tones were also observed (e.g., subject 1 at 50 Hz). For f_s = 31.5 and 40 Hz, addition of the MDI tones did not affect the trend for the tips of the PTCs to fall above f_s. For the mean PTCs obtained with MDI tones, the steepness in dB/oct of the lower skirts was about 25, 23, 25, 22, and 25 dB/oct, for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively. The data were not sufficient to determine the steepness of the upper skirts.

It should be noted that, while the MDI tones seem to have been effective in reducing the effectiveness of beats as a cue, the presence of the MDI tones generally made the task subjectively more difficult. Furthermore, the possibility cannot be ruled out that the MDI tones had a direct masking effect on the signal in some cases. The origin of the PTCs with sharp tips in the presence of the MDI tones, such as that for subject 1 for f_s = 50 Hz, is not clear. They may reflect individual differences in susceptibility to MDI; the magnitude of MDI can vary across subjects, and sometimes it almost disappears after extensive training (Bacon and Opie, 1994).

For cases where PTCs were also measured with the noise masker (see dots

and dotted lines in figures 3 and 4; subjects 2, 4 and 7), the PTCs had similar overall features to the PTCs derived with the sinusoidal masker alone. However, the PTCs for the noise masker were more regular around their tips, resembling the PTCs derived with MDI tones. The PTCs derived with the noise maskers are described in more detail in sub-section F.

D. Values of the ERB and 10-dB bandwidth using sinusoidal maskers without and with MDI tones

The mean PTCs obtained using sinusoidal maskers without and with MDI tones are shown in Fig. 5(A) (circles and asterisks, respectively). Table I shows ERBs and 10-dB bandwidths calculated from these PTCs ('Tones' and '+MDI' columns, respectively). The locations of the tips of the mean PTCs are also given in the table. To calculate the bandwidths of the PTCs obtained with MDI tones, masker levels at threshold for values of Δ outside the region where the MDI tones were used were taken as those obtained without MDI tones. For the PTCs obtained without MDI tones the ERB varied markedly with f_s, being smallest at f_s = 80 Hz, but not showing a clear trend with frequency below that. The much smaller value at f_s = 80 Hz occurred because most of the PTCs for f_s = 80 Hz had tips at f_s and had sharp tips, presumably due to the influence of beats; beats provide a salient detection cue for masker frequencies adjacent to f_s, but do not occur when $f_m = f_s$.

The mean ERBs calculated from analysis of the individual data were 34.8, 38.0, 32.3, 38.5, and 22.7 Hz, for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively, with standard deviations (SDs) of 10, 9.5, 7.9, 9.5, and 11.6 Hz. For f_s = 80 Hz the ERB value was smallest, and was similar to that of the mean PTC, while for f_s between 40 and 63 Hz the ERB values were smaller than for the mean PTCs. This happened mostly because the frequency of the tips of the PTCs varied across subjects, resulting in broadening of the tips of the mean PTCs.

The values of the 10-dB bandwidths were larger by a factor of 1.5 or more than the ERB values. When the 10-dB bandwidths were calculated using f_s as the reference point, the bandwidths were always larger than when they were calculated using the tip frequency as the reference point. This happened because, when the tip was above f_s, the skirts were less sharp around f_s than around the tip, and because, when bandwidths were calculated using f_s as the reference point, the region of decreasing masker level between f_s and the tip frequency was included in the calculation of the width of the upper side. For the PTCs obtained without MDI tones, the 10-dB bandwidths expressed relative to f_s, tended to increase with decreasing f_s.

For the mean PTCs obtained with MDI tones, the ERB values decreased with decreasing f_s down to f_s = 40 Hz, but increased at f_s = 31.5 Hz. The largest difference in ERB values between the PTCs occurred for f_s = 80 Hz, where the ERB roughly doubled when the MDI tones were added. This happened because most of the individual PTCs for f_s = 80 Hz had a sharp tip at f_s in the absence of the MDI tones, and the tips were broadened when the MDI tones were present. The MDI tones also led to an increase in the ERB for f_s = 31.5 Hz, but for f_s between 40 and 63 Hz, the ERB values were smaller with MDI tones than without them. This appears to have happened because addition of the MDI tones removed some of the irregularities in the PTCs around the tips, producing more consistent and more distinct tips. The 10-dB bandwidths of the PTCs obtained with MDI tones did not show a clear trend with f_s.

Table I. Equivalent rectangular bandwidths (ERB), 10-dB bandwidths (10-dB BW), asymmetry (A_{10dB}) and locations of the tips (Tip frequency), derived from the mean PTCs obtained with the sinusoidal maskers, sinusoidal maskers with MDI tones, and narrowband-noise maskers ('Tone', '+MDI', and 'Noise' columns, respectively). Asymmetry and 10-dB bandwidths are expressed both relative to the tip and to f_s (denoted tip| f_s). Only one value is given when the tip of the PTC was located at f_s. Values indicated [a] are based on extrapolation. For f_s = 63 Hz, the tip was broad, and the masker level at 63 Hz was only 0.1 dB above the level at 88 Hz.

	Tip frequency (Hz)			ERB (Hz)			10-dB BW (Hz)			$A_{10\,dB}$		
f_s	Tone	+MDI	Noise	Tone	+MDI	Noise	Tone tip|f_s	+MDI tip|f_s	Noise tip|f_s	Tone tip|f_s	+MDI tip|f_s	Noise tip|f_s
80	80	96	80	20.7	43.2	31.7	51	69.1|71.5	57.9	0.4	1.9|0.7	0.8
63	88	76	63	51.9	30.0	34.3	84.5|84.9	69.8|85	61.4	0.8|0.2	0.9|0.4	0.5
50	70	60	60	37.9	29.7	39.7	73.8|103.1	64.8|80.9	70.6|74.6	0.6|0.3	0.6|0.3	0.5|0.3
40	88	56	48	46.8	25.8	51.9	70.8|99.2	56.8|82	87.7|96.4	2.3|0.1	0.5|0.2	0.3|0.1
31.5	69	44	69	33.6	45.4	53.3	69 |124.9[a]	66.2|94.6	86.0|110.8[a]	0.8|0.1[a]	0.3|0.1	0.9|0.1[a]

E. Asymmetry of PTCs obtained using sinusoidal maskers without and with MDI tones

The values of A_{10} dB are shown in the right-most column of Table I, in the two columns labeled 'Tone' and '+MDI'. These were calculated from the mean PTCs shown in Fig. 5(A). The values of A_{10} dB were always smaller when expressed relative to f_s than relative to the tip frequency. Expressed relative to f_s, A_{10} dB tended to decrease with decreasing f_s, indicating that the lower skirt of the PTC became progressively sharper than the upper skirt as f_s decreased. Analysis based on the individual data revealed a similar trend. The relative standard deviations (RSD, SD expressed as a percentage of the mean) in asymmetry obtained for the PTCs measured with MDI tones were 28.3, 41.9, 36.3, 24.1, and 29% for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively. For the PTCs obtained without MDI tones, the values were 61.4, 51.1, 53.2, 24.7, and 49.9%, respectively. The somewhat larger RSD values found at low f_s are probably due to larger individual differences in the sharpness of the upper skirts, as can be observed from the individual data in Figs. 3 and 4.

The overall trends in asymmetry are consistent with those observed for f_s down to 50 Hz for auditory filter shapes derived from notched-noise masking data by Jurado and Moore (2010), and are the opposite of what is typically observed at higher frequencies.

F. PTCs obtained using the noise masker

The mean PTCs obtained with the noise masker are shown in Fig. 5(B); individual PTCs are shown in Figs. 6 and 7. The PTCs generally had more regular tips than those obtained using the sinusoidal masker (without MDI tones), suggesting that beats probably did not provide a strong detection cue for the noise masker. The tips of the PTCs were somewhat broader than those found for the sinusoidal masker with MDI tones. Even though the masker bandwidth was proportional to f_s, the PTCs broadened (on a log-frequency scale) as f_s decreased.

As was the case for the PTCs obtained using the sinusoidal masker, in all cases the PTCs for f_s = 31.5 and 40 Hz had tips that fell above f_s. For f_s = 50, 63, and 80 Hz, the tips were at f_s in 2, 5, and 7 out of 8 cases,

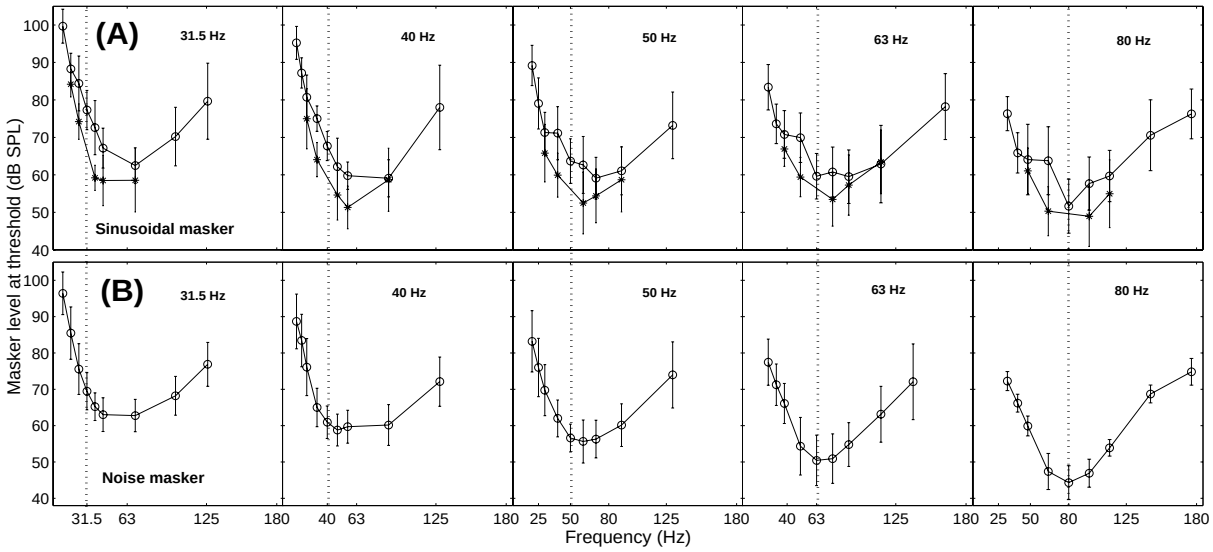

Figure 5. (A). Mean PTCs obtained with the sinusoidal masker without MDI tones (open circles) and with MDI tones (asterisks). (B) Mean PTCs obtained with the noise maskers. The values of f_s are shown in each panel. Error bars indicate ±1 SD. Vertical dotted lines indicate the value of f_s.

respectively, and otherwise fell above f_s. Again, the shifted tips for the lowest values of f_s may indicate that there are no auditory filters tuned below about 50 Hz, as suggested by Moore et al. (1997). However, upward-shifted tips of the PTCs may also be produced by the steep slope of the METF in this region (Jurado et al., 2010). A detailed analysis of the estimated influence of the METF on the PTCs is given in section IV.

The masker levels at threshold tended to be higher for lower values of f_s. This was also the case for the PTCs obtained with the sinusoidal maskers, and is probably mainly a consequence of the steep increase in absolute thresholds with decreasing frequency in this frequency range. Mean absolute thresholds were 53.4, 43.1, 39.9, 32.5, and 26.3 dB SPL for f_s = 31.5, 40, 50, 53, and 80 Hz, respectively. On average, the maximum and minimum thresholds increased with decreasing f_s, while the range of thresholds was roughly similar across f_s. The (max, min, range) values were (96.4, 62.8, 45.6), (88.7, 58.8, 42.9), (83.2, 55.6, 41.6), (77.5, 50.4, 42.1), and (74.8, 44.3, 46.5) dB, for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively.

The masker level required to mask the signal around the tips of the PTCs was somewhat lower than for the sinusoidal masker, probably due to the inherent amplitude fluctuations in the noise masker, which introduce level uncertainty and make beats less effective as a detection cue. The average difference between the signal and masker levels, when the masker was centered at f_s, was –1.0, –2.8, –2.9, –3.0, and –3.0 dB for f_s = 31.5, 40, 50, 53, and 80 Hz, respectively. These values are close to the values of K, the estimated the signal-to-noise ratio at the output of the auditory filter required to achieve the detection threshold, derived by Jurado and Moore (2010) from notched-noise masking data. In their "System-as-a-whole" analysis, comparable to the present results, they obtained values of –2.1, –2.2, and –0.2 for f_s = 50, 63, and 80 Hz, respectively. The differences are also similar, but somewhat smaller, than those obtained by Kluk and Moore (2004) for

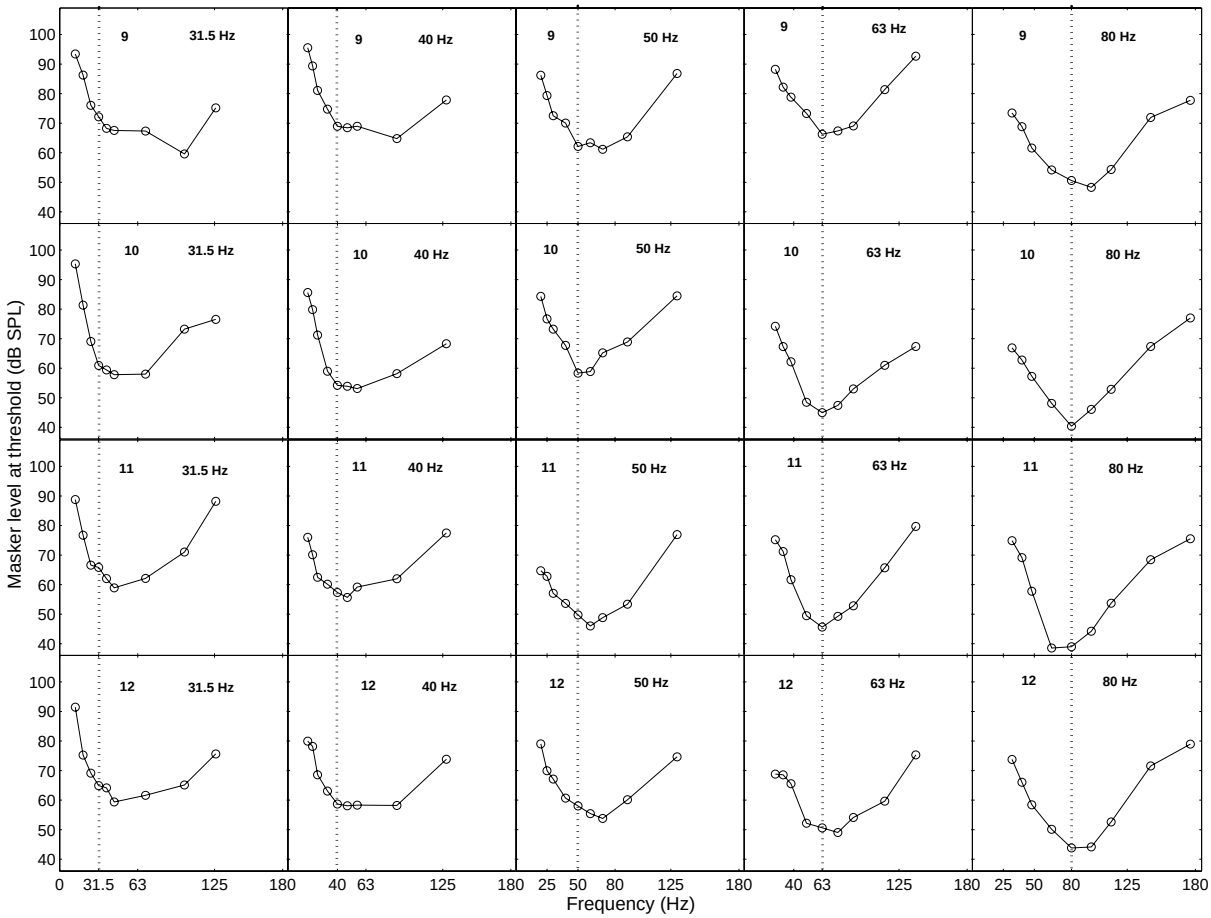

Figure 6. Individual PTCs for four subjects obtained using the noise masker. Otherwise, as Fig. 3.

PTCs measured using higher signal frequencies.

For the mean PTCs, the slopes in dB/oct of the lower and upper skirts (lower, upper) were about (21, 33), (23, 30), (21, 32), (23, 31), (25, 35) for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively. The ratio of slopes of the upper and lower skirts was, on average, about 1.4, the same as for the sinusoidal masker. On a linear frequency scale, the PTCs became more asymmetric with decreasing f_s, with steeper lower skirts than upper skirts; again this is the same as found for the sinusoidal masker. There were clear individual differences. For example, for subject 9 the upper skirt of the PTC for f_s = 31.5 Hz was very shallow, only starting to increase for the highest masker frequency used, while for subject 11 the upper skirt started increasing for frequencies much closer to f_s. Individual differences are discussed in more detail below.

G. Values of the ERB and 10-dB bandwidth for PTCs obtained with the noise masker

The values of the ERB and 10-dB bandwidth obtained from the mean PTCs shown in Fig. 5(B) are given in Table I ('Noise' columns). The locations of the tips are also given. There was a clear tendency for the bandwidths (including 10-dB bandwidths) to increase with decreasing f_s, reflecting the worsening of tuning with decreasing frequency in the frequency range between f_s = 31.5 and 80 Hz.

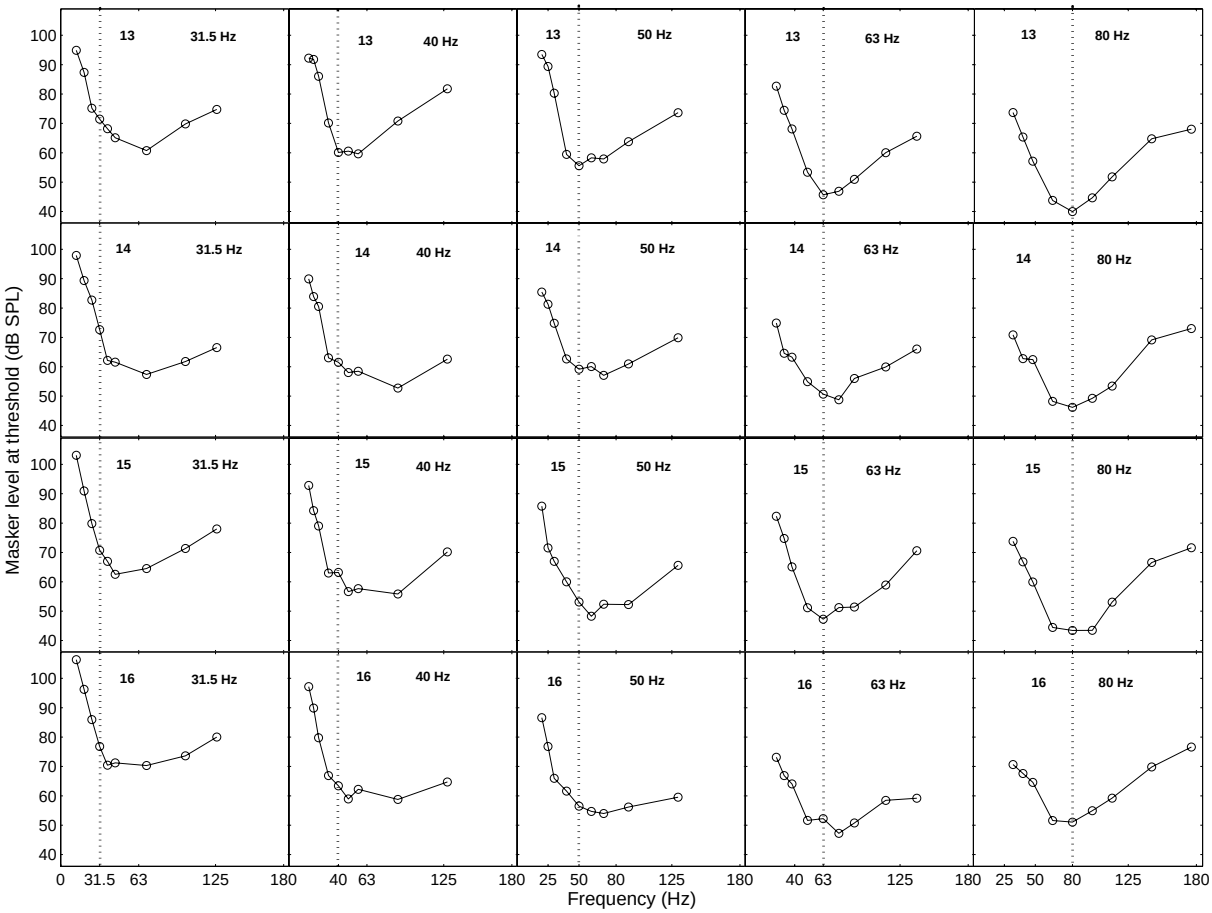

Figure 7. As Fig. 6, but for four more subjects.

The mean ERB values derived from the individual PTCs were 40.4, 44.0, 32.9, 28.3, and 31.6 Hz for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively. These ERB values are somewhat smaller than the ERB values derived from the mean PTCs, due to tip frequency varying across individuals. The SDs of the ERB values obtained from the individual PTCs were 10.1, 11.4, 9.6, 2.8, and 6.4 Hz for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively. The higher SDs for the lowest values of f_s are probably a product of individual differences in sharpness of the upper skirts (see sub-section H).

The ERB values derived from the mean PTCs for the noise masker were higher than those obtained using the sinusoidal masker with MDI tones for all signal frequencies except 80 Hz. The ERB values derived from the mean PTCs for the noise masker are in agreement with ERB values estimated by Jurado and Moore (2010) from notched-noise masking data for f_s down to 50 Hz. For their "System-as-a-whole" analysis (which treated the METF as part of the auditory filter, and is comparable to the "raw" PTCs considered so far here) their ERB values were 37.6, 30.9, and 24 Hz for f_s = 50, 63, and 80 Hz, respectively.

H. Asymmetry of the PTCs obtained with noise maskers

The values of A_{10} dB are given in the right-most column of Table I. A_{10} dB was always below 1, indicating that the lower skirts of the PTCs were sharper than the upper skirts for f_s in the range 31.5 to 80 Hz. As for the PTCs obtained with the sinusoidal

masker, A_{10} dB clearly decreased with decreasing f_s, when calculated relative to f_s, reflecting the progressive sharpening of the lower skirt and simultaneous flattening of the upper skirt with decreasing f_s. For f_s = 31.5 and 40 Hz, the upper skirt was very shallow in the frequency region above f_s. These shallow upper skirts are at least partly a consequence of the effect of the METF; this is considered in more detail later on. The RSD values of A_{10} dB, calculated relative to f_s for the individual data, were 45.7, 42.6, 36.0, 28.3, and 15.2% for f_s = 31.5, 40, 50, 63, and 80 Hz, respectively. The values increased with decreasing f_s, reflecting individual differences in the sharpness of the upper skirts and differences in the locations of the tip of the PTCs.

The trends in asymmetry and their values are in good agreement with results obtained by Jurado and Moore (2010) from notched-noise masking data for f_s down to 50 Hz (analyzed for the "System-as-a-whole" case), where asymmetry was found to reverse from steeper upper skirts above about 100 Hz to steeper lower skirts below that. Their asymmetry index values were 0.3, 0.5, and 0.9 for f_s = 50, 63, and 80 Hz, respectively, which are roughly the same as obtained here. Such reverse behavior from what is typically found at higher frequencies has been observed in physiological measurements (Cheatham and Dallos, 2001) and is assumed to be due to highpass-filter action of the METF at low frequencies.

I. Equal-loudness contours

The ELCs for the 14 individual subjects (thin solid lines) and the mean across subjects (thick dashed line) are shown in Fig. 8. The ELCs appear to have three regions: (1) A region above about 100 Hz with a shallow slope; (2) A transition region, more irregular in shape (covering a range from about 35 to 100 Hz in the mean data); and (3) A steep region below about 35 Hz. For the mean data, the slopes in regions 1 and 3 were about −10 dB/oct and −21 dB/oct, respectively. The three regions are in qualitative agreement with objective measures of the METF reported by Marquardt and Pedersen (2010). According to models of the low-frequency characteristics of the METF (Marquardt et al., 2007; Marquardt and Hensel, 2008) : the slope in region (1) should be about −6 dB/octave, reflecting the effect of middle-ear stiffness; region (2) is a transition region that reflects a cochlear resonance (roughly, a "hump") produced by the impedance of the helicotrema; and the slope in region (3) should be about −12 dB/oct, being dominated by the shunt effect of the helictrema. The slopes of the ELCs are markedly steeper than the slopes of the METF reported by Marquardt and Pedersen (2010).

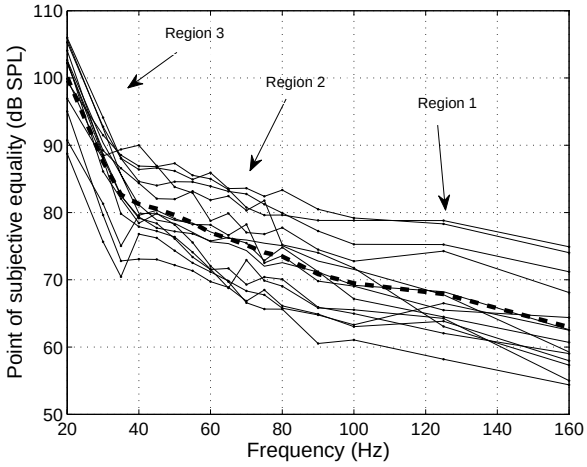

Figure 8. Individual (solid lines) and mean (dashed line) equal-loudness contours. The values at 50 Hz correspond to a level of 40 dB SL. The three regions observed are indicated by the arrows.

The difference may partly arise from the dependence of neural spike rates on the velocity of inner hair cells (Cheatham and Dallos, 2001), since the ELCs presumably depend on neural spike rates rather than

depending directly on basilar-membrane velocity. The difference may also arise because the ELC provides only an approximate estimate of the METF (Moore et al., 1997).

The individual ELCs reported here show similar irregularities to those reported by Marquardt et al. (2007) and Marquardt and Pedersen(2010), including a "hump" for some subjects. However, the mean ELC is rather smooth, as the exact frequency span and center frequency of the hump varied across subjects. A more detailed description of the individual ELCs, together with a comparison of these with objective estimates of the subject's individual METFs obtained through a distortion-product-isomodulation technique is given in (Jurado, 2010). The mean ELCs of the two group of subjects tested using the sinusoidal and noise maskers (both groups composed of 8 subjects) had similar overall features to the mean ELC shown in Fig. 8, although region 3 (the steeper region) started at about 40 Hz for the group of subjects tested using noise.

IV. INFLUENCE OF THE METF ON TUNING CHARACTERISTICS

In principle, fixed filtering processes applied to sounds before they reach the cochlea, such as the transmission of sound through the middle ear, can be treated as separate from the processes in the cochlea (and perhaps higher levels in the auditory system) that are associated with frequency selectivity (Moore and Glasberg, 1987; Glasberg and Moore, 1990). The PTCs presented so far reflect the combined effects of processes at all levels in the auditory system. In this section, we consider the tuning curves that result from treating the METF as a fixed filter occurring prior to auditory frequency analysis. This method of analyzing the results is comparable to accepted methods for analyzing notched-noise data (Glasberg and Moore, 1990; Unoki et al., 2006), and is comparable to the "METF-separate" analysis described by Jurado and Moore (2010). It should be noted, however, that the METF includes the shunt effect of the helicotrema, which has a strong influence below about 40 Hz and which is a cochlear mechanism (Marquardt and Hensel, 2008; Marquardt and Pedersen, 2010), not a middle-ear mechanism. Therefore, in what follows, special emphasis is placed on the characteristics of tuning in the region below about 40 Hz.

For simplicity, the inverse of the ELC curve will be taken as corresponding to the METF (Moore and Glasberg, 1987). The tuning curves obtained after allowing for the effects of the METF can be thought of as analogous to inverted auditory filter shapes. For brevity, these will be referred to simply as TCs. In the analysis presented below, we focus on the results obtained using the noise masker, as the results obtained with the sinusoidal masker were probably influenced by detection of beats, and the results obtained with the sinusoidal masker plus MDI tones may have been influenced by direct masking effects of the MDI tones.

Figure 9 shows the TCs obtained after allowing for the METF, obtained from the mean PTCs for the noise masker. The left panel shows results for f_s = 31.5 and 40 Hz and the right panel shows results for f_s = 50, 63, and 80 Hz. To make the TCs resemble filter characteristics, they have been inverted and normalized by plotting the level at the tip at 0 dB. The METF is plotted with an arbitrary reference.

The PTCs had tips at 69.3, 48, and 60 Hz, for f_s = 31.5, 40, and 50 Hz, respectively. The TCs for f_s = 31.5 and 40 Hz did not have well defined tips, but had more of a lowpass characteristic, with edge frequencies of 44 and 48 Hz, respectively. The TC for f_s = 50 Hz had a tip at 50 Hz.

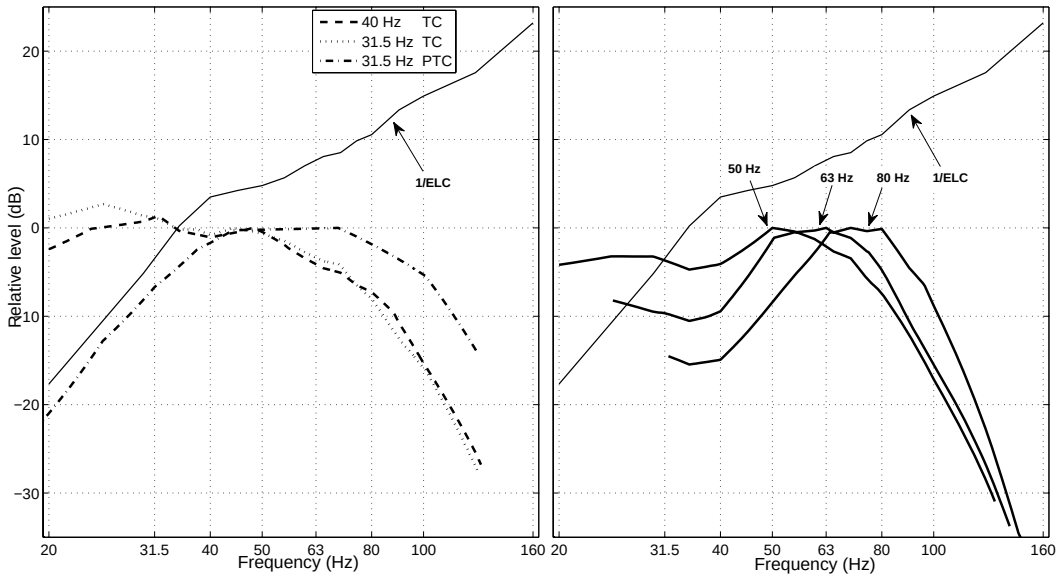

Figure 9. TCs obtained after allowing for the METF for PTCs obtained with a noise masker. The mean (inverted) ELC is plotted as a thin-solid line in each panel at arbitrary level. The left panel shows TCs for f_s = 31.5 and 40 Hz (dotted and dashed lines, respectively) together with the PTC for f_s = 31.5 (thin dashed-dotted line) for comparison; the right panel shows the TCs for f_s = 50, 63, and 80 Hz.

These results may indicate that the "bottom"auditory filter has a CF a little below 50 Hz. However, it is possible that the "true" lowest CF is slightly below the lowest values obtained. This is so because the masker had a bandwidth of 0.5f_s, and when the center frequency of the masker coincided with the CF of the lowest auditory filter, part of the masker spectrum would have fallen in the steep region of the METF below 40 Hz, and would have been strongly attenuated. Thus, higher masker levels would be required to mask the signal, leading to a TC with a tip slightly above the lowest CF.

Results of Jurado et al. (2010) indicate that if an auditory filter was assumed to be centered near the start of region 3 in the METF (i.e around 40 Hz here), the "combined" filter shape representing the frequency selectivity of the "System as a whole" typically had a tip which was also close to this point. This was different for auditory filters assumed to be centered within region 2, where the tips of the combined filters were generally shifted upwards in frequency relative to f_s as a result of the steepness of the METFs in that region. In the present work, for f_s = 40 Hz, the tip of both the PTC and the TC were at the same frequency, at 48 Hz, which is close to but slightly above f_s. This does not conflict with the notion that the most apical auditory filter may be centered close to the start of region 3, because the finite masker bandwidth may have shifted the tip slightly above the frequency corresponding to the lowest CF, for the reasons described above.

The lower skirts of the TCs flattened off considerably below the point where the METF increases in steepness, i.e. below about 40 Hz. The effect was particularly marked for f_s = 31.5 and 40 Hz, since the PTCs for these values of f_s had lower skirts with similar slopes to the METF (about 21 and 23 dB/oct, respectively). Hence, allowing for the effect of the METF resulted

Table II. Equivalent rectangular bandwidths (ERB), 10-dB bandwidths (10-dB BW), and corresponding sharpness of tuning (Q_{ERB} and Q_{10dB}) for the mean TCs obtained using the noise masker by treating the METF as a fixed filter prior to auditory filtering [case (a)] and by including the METF as part of frequency selectivity only for frequencies below 40 Hz [case (b); composite TCs]. No values of the 10-dB bandwidth or Q_{10dB} are given for case (a) for values of f_s where there was not a 10-dB range on the lower skirt. The locations of the tip frequencies of the TCs are also given. For f_s = 80 Hz, the tip was broad, and the masker level at 80 Hz was only 0.1 dB above the level at 70 Hz.

	Tip frequency	ERB (Hz)		10-dB BW (Hz)		Q_{ERB}		Q_{10dB}	
f_s	(Hz)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
80	70	31.5	31.4	54.6	54.6	2.5	2.6	1.5	1.5
63	63	30.6	29.5	51.0	49.6	2.1	2.1	1.2	1.3
50	50	33.8	27.5	--	53.2	1.5	1.8	--	0.9
40	48	41.9	28.6	--	60.2	1.0	1.4	--	0.7
31.5	44	44.3	30.1	--	57.2	0.7	1.0	--	0.6

in low-frequency sides of the TCs that were almost flat. This suggests that the region of the METF below about 40 Hz, corresponding to the region where the shunt effect of the helicotrema is dominant, is an inherent part of tuning that largely defines the low-frequency side of the most apical auditory filters.

The use of inverted ELC curves as approximation of the METF may lead to a correct analysis of the slopes of the low-frequency sides of the auditory filters at very low frequencies. However, it is also possible that the slope of the METF is less than estimated from the ELC, in which case the "true" TCs would have finite slopes on their low-frequency sides.

The TCs were very similar for f_s = 31.5 Hz and 40 Hz. This is consistent with the idea that the signals for both of these values of f_s were detected using the same auditory filter, a filter with an edge or tip frequency located between 40 and 50 Hz.

Analyses similar to those described above were conducted on the individual PTCs for the noise masker, using the individual estimates of the METF. The results were consistent with those for the mean data, although there were individual differences in the inferred location of the most apical CF, which varied from 40 to 65 Hz and typically lay between 40 and 50 Hz.

The ERBs and 10-dB bandwidths of the TCs are shown in Table II, under the headings (a) in columns three and four. The numbers under heading (b) show results of an additional analysis, in which the part of the METF below 40 Hz was treated as a part of frequency selectivity; this seems reasonable, as that part of the METF is dominated by the shunt action of the helicotrema, which is a cochlear process rather than a middle-ear process. We refer to tuning curves derived in this way as "composite TCs". Below 40 Hz, the composite TCs have the same shape as the PTCs (without allowing for the METF), while above 40 Hz the composite TCs have the same shape as the TCs.

The bandwidths of the TCs [case (a)] tended to increase when f_s decreased below 63 Hz. This is consistent with the auditory filter shapes derived by Jurado and Moore (2010) from analysis based on notched-noise masking data (measured down to f_s = 50 Hz) and is a consequence of the progressive flattening of the lower skirts of the TCs as f_s decreases. In contrast, the bandwidths of the composite TCs [case (b)] were roughly independent of f_s. The ERB values were

close to the value of 30 Hz predicted for low CFs from the formula given by Glasberg and Moore (1990). The tendency for the bandwidth to reach an asymptote for very low CFs is consistent with the comparable analysis of Jurado and Moore (2010) of notched-noise masking data, in which the METF was treated as part of the auditory filter shape for frequencies below 50 Hz.

The values of Q_{ERB} and Q_{10} dB, which are measures of the sharpness of tuning, are also given for cases (a) and (b) in Table II. The values are much smaller than the values of about 7.6 and 4, respectively, for CF = 1000 Hz (calculated from Glasberg and Moore, 1990). The relative sharpness of tuning clearly worsens at low frequencies (it is about 3 times poorer at 80 Hz than at 1000 Hz).

V. CONCLUSIONS

We have measured psychophysical tuning curves for values of f_s between 31.5 and 80 Hz, using sinusoidal maskers, sinusoidal maskers plus beating (MDI) tones, and narrowband-noise maskers. The following are our main conclusions:

(1) The PTCs for the low signal frequencies used here were broad, indicating that frequency selectivity is relatively poor at low frequencies. For example at f_s = 80 Hz, the relative sharpness of tuning was about 1/3 of that typically observed at 1000 Hz.

(2) For the noise maskers, the bandwidths of the PTCs showed a tendency to increase with decreasing f_s below 80 Hz, with larger standard deviations at the lowest values of f_s.

(3) The tips of the PTCs for f_s = 31.5 and 40 Hz always fell above f_s. As f_s increased, the tips fell at f_s more often, and were typically at f_s for f_s = 80 Hz.

(4) The PTCs became increasingly asymmetrical with decreasing f_s, with steep lower skirts and shallow upper skirts (on a linear frequency scale).

(5) Individual differences in asymmetry generally increased as f_s decreased, mostly due to individual variation in the sharpness of the upper skirts.

(6) The PTCs obtained with the sinusoidal maskers without MDI tones were generally more irregular around their tips than those measured with MDI tones or measured using noise maskers. Comparison of PTCs for the sinusoidal masker alone and the sinusoidal masker with MDI tones suggested that beats between the signal and masker provided a powerful detection cue in the former case for relative frequency separations, Δ, up to 0.4. The effect of the MDI tones was marginal for Δ = 0.8.

(7) The shapes of the equal-loudness contours were in qualitative agreement with the shapes of METFs derived from a distortion-product-isomodulation technique by Marquardt and Pedersen (2010), but the former had somewhat steeper slopes.

(8) Treating the METF as a fixed filter applied before auditory filtering resulted in TCs with steeper upper skirts, reduced the frequencies at the tips of the PTCs for f_s between 31.5 and 50 Hz, and flattened the lower skirts, especially below 40 Hz.

(9) The region in the METF below about 40 Hz is dominated by the helicotrema shunt mechanism. It may be reasonable to treat this as an inherent part of frequency selectivity, largely defining the low-frequency skirts of the most apical auditory filters.

(10) The CF of the "bottom" auditory filter (the one with the most apical CF) appears to be located between 40 and 50 Hz.

ACKNOWLEDGMENTS

We thank Claus Vestergaard and Peter Dissing for their help with the experimental setup. Additionally we thank Steffen Stengård Villadsen for his assistance in carrying out measurements validating the setup. We are also deeply thankful to all participants.

REFERENCES

Alcántara, J. I., Moore, B. C. J., and Vickers, D. A. (**2000**). "The relative role of beats and combination tones in determining the shapes of masking patterns at 2 kHz: I. Normal-hearing listeners," Hear. Res. **148**, 63-73.

Bacon, S. P., and Moore, B. C. J. (**1993**). "Modulation detection interference: Some spectral effects," J. Acoust. Soc. Am. **93**, 3442-3453.

Bacon, S. P., and Opie, J. M. (**1994**). "Monotic and dichotic modulation detection interference in practiced and unpracticed subjects," J. Acoust. Soc. Am. **95**, 2637-2641.

Cheatham, M. A., and Dallos, P. (**2001**). "Inner hair cell response patterns: implications for low-frequency hearing," J. Acoust. Soc. Am. **110**, 2034-2044.

Egan, J. P., and Hake, H. W. (**1950**). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. **22**, 622-630.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103-138.

Glasberg, B. R., Moore, B. C. J., Patterson, R. D., and Nimmo-Smith, I. (**1984**). "Dynamic range and asymmetry of the auditory filter," J. Acoust. Soc. Am. **76**, 419-427.

Greenwood, D. D. (**1971**). "Aural combination tones and auditory masking," J. Acoust. Soc. Am. **50**, 502-543.

ISO 226 (**2003**). *Acoustics - normal equal-loudness contours* (International Organization for Standardization, Geneva).

Jurado, C. A. (**2010**). "Behavioral estimates of human frequency selectivity at low frequencies," Ph.D. Thesis, Aalborg University.

Jurado, C. A., and Moore, B. C. J. (**2010**). "Frequency selectivity for frequencies below 100 Hz: comparisons with mid-frequencies," J. Acoust. Soc. Am. (in press).

Jurado, C. A., Pedersen, C. S., and Marquardt, T. (**2010**). "Frequency selectivity at very low centre frequencies: the influence of the helicotrema on individual differences in low-frequency sound perception," in *Proceedings of 14th International Conference on Low Frequency Noise and Vibration and its Control* edited by F. Christensen, and C. S. Pedersen (Aalborg, Denmark), pp. 129-145.

Kluk, K., and Moore, B. C. J. (**2004**). "Factors affecting psychophysical tuning curves for normally hearing subjects," Hear. Res. **194**, 118-134.

Leventhall, H. G. (**2004**). "Low frequency noise and annoyance," Noise and Health **6**, 59-72.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467-477.

Marquardt, T., and Hensel, J. (**2008**). "A lumped-element model of the apical cochlea at low frequencies," in *Concepts and Challenges into the Biophysics of Hearing*, edited by N. P. Cooper, and D. T. Kemp (World Scientific Publishing Co, London), pp. 337-339.

Marquardt, T., and Pedersen, C. S. (**2010**). "The infuence of the helicotrema on low-frequency hearing," in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York), pp. 25-36.

Marquardt, T., Hensel, J., Mrowinski, D., and Scholz, G. (**2007**). "Low-frequency characteristics of human and guinea pig cochleae," J. Acoust. Soc. Am. **121**, 3628-3638.

Møller, H., and Pedersen, C. S. (**2004**). "Hearing at low and infrasonic frequencies," Noise Health **6**, 37-57.

Moore, B. C. J., and Glasberg, B. R. (**1987**). "Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns," Hear. Res. **28**, 209-225.

Moore, B. C. J., Alcántara, J. I., and Dau, T. (**1998**). "Masking patterns for sinusoidal and narrowband noise maskers," J. Acoust. Soc. Am. **104**, 1023-1038.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). "A model for the prediction of thresholds, loudness and partial loudness," J. Audio Eng. Soc. **45**, 224-240.

Patterson, R. D. (**1976**). "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. **59**, 640-654.

Patterson, R. D., and Nimmo-Smith, I. (**1980**). "Off-frequency listening and auditory filter asymmetry," J. Acoust. Soc. Am. **67**, 229-245.

Unoki, M., Irino, T., Glasberg, B. R., Moore, B. C. J., and Patterson, R. D. (**2006**). "Comparison of the roex and gammachirp filters as representations of the auditory filter," J. Acoust. Soc. Am. **120**, 1474-1492.

Vogten, L. L. (**1978**). "Low-level pure-tone masking: a comparison of "tuning curves" obtained with simultaneous and forward masking," J. Acoust. Soc. Am. **63**, 1520-1527.

Vogten, L. L. M. (**1974**). "Pure-tone masking: A new result from a new method," in *Facts and Models in Hearing*, edited by E. Zwicker, and E. Terhardt (Springer-Verlag, Berlin), pp. 142-155.

Wegel, R. L., and Lane, C. E. (**1924**). "The auditory masking of one sound by another and its probable relation to the dynamics of the inner ear," Phys. Rev. **23**, 266-285.

Yost, W. A., and Sheft, S. (**1989**). "Across-critical-band processing of amplitude-modulated tones," J. Acoust. Soc. Am. **85**, 848-857.

14th International Meeting on
Low Frequency Noise and Vibration and its Control
Aalborg, Denmark 9 – 11 June 2010

Frequency selectivity at very low centre frequencies: the influence of the helicotrema on individual differences in low-frequency sound perception.

Carlos A. Jurado[a)]

Christian S. Pedersen
Section of Acoustics, Aalborg University, Fredrik Bajers Vej 7 B5 , DK-9220 Aalborg, Denmark
[a)] **E-mail: cjo@es.aau.dk**

Torsten Marquardt
UCL Ear Institute, 332 Grays Inn Rd London, WC1X 8EE, UK

Summary

A significant factor in the decrease of sensitivity to low-frequency sound is the helicotrema shunt effect. In humans, it causes a slope increase of the middle-ear transfer function (METF) from 6 dB/oct to 12 dB/oct below approximately 50 Hz [Marquardt *et al*., J.Acoust. Soc. Am. 121, 3628-3638 (2007)]. Recent experiments showed that the exact frequency varies from individual to individual. Besides, the helicotrema region in the METF has been found to highly influence frequency selectivity for centre frequencies (CFs) below 80 Hz (Jurado and Moore in prep). By using individual METF measurements based on OAE suppression techniques and notched-noise masking data psychophysically measured for centre frequencies in the range 50-125 Hz, this study examines how individual differences in frequency selectivity, as well as in masking, may occur at very low CFs due to individual differences in the shape of the METFs, thought to be affected by the helicotrema impedance. Preliminary analysis indicates that individual differences in the METFs might underlay the observed individual differences in frequency selectivity. Main effects predicted by the calculations are a pronounced flattening off of the upper filter skirts and a shift in the tip, typically upwards in frequency, for auditory filters with CFs below about 80 Hz.

1. Introduction

Many environmental sounds contain substantial energy at low frequencies. Due to their large wavelengths, such sounds can easily propagate and produce annoyance (Leventhall 2004). While audibility and loudness of pure low-frequency tones can be predicted by the use of standardized threshold curves and equal-loudness contours (ISO 226 2003), for complex low-frequency signals such predictions need to be based on auditory models, which depend on

information about the auditory filter characteristics (see e.g. Moore et al. (1997)). Only recently, Jurado and Moore (in prep) have obtained frequency selectivity estimates for center frequencies (CFs) between 50 Hz and 1000 Hz, thus covering the "unexplored" range of frequencies below 100 Hz Their results in terms of the bandwidth of the auditory filter were in broad agreement (down to 80 Hz) with the ERB function suggested by Glasberg and Moore (1990), that indicates the ERB continues to decrease with decreasing CF below 500 Hz. This, as many other evidence (see e.g. Moore et al. 1990; Rosen and Stock 1992; Unoki et al. 2006), contradicts the notion that the ERB flattens off reaching a value of about 100 Hz below 500 Hz, as suggested by Zwicker et al. (1957).

The new results from Jurado and Moore (in prep) indicate also that the 12 dB/oct region in the middle-ear transfer function (METF) highly influences frequency selectivity CFs below 80 Hz. Excluding this region of effective helicotrema shunt effect (located below about 50 Hz in the METF), produced very shallow lower filter skirts (and therefore larger bandwidth values), indicating below about this frequency the helicotrema shunt accounts for most of frequency selectivity. On the other hand, frequency selectivity of the system as a whole (i.e. including the effect of the middle-ear transfer function) is relatively poor for the CFs of 63 and 50 Hz, where the equivalent rectangular bandwidth (ERB) increased with decreasing CF, in great part due very shallow upper filter skirts observed in the derived auditory filer shapes. However, their main analysis was focused on mean masking data across subjects, where a "default" METF was used as an approximation of an average attenuation curve (based on the work of Marquardt et al. (2007) and Marquardt and Hensel (2008)), and individual differences in frequency selectivity were not treated in that work. Regarding hearing thresholds and equal loudness contours, there is evidence of microstructures in the low-frequency characteristics of these curves, which may be related to individual differences in perception of low frequency sound and reports of "especially sensitive" persons (see Møller and Pedersen (2004) for a review).

Nevertheless, it is the shape of the *auditory filter* what will finally determine the excitation patterns internally created by (complex) low-frequency sound, and thus influence the perception of such acoustic signals. This evinces the importance of describing frequency selectivity in the low-frequency region, as well as describing the degree of individual differences and understanding the possible factors affecting those. Recently, Marquardt and Pedersen (2010) estimated individual shapes of the METF objectively by the use of distortion product isomodulation techniques (DPOAE). Their results suggest that there is an irregular frequency region in the METF which separates an approx 6 dB/oct slope region (above it) and an approx 12 dB/oct region (below it). The extent of this region is rather individual, and consists of a small depression (or elevation considering 1/METF). Individual differences in the frequencies of the dip in the depression (up to 25 Hz in their 5 subjects) were attributed as possible differences in helicotrema size and shape.

Due to the importance of the helicotrema region in the METF in determining frequency selectivity at very low CFs, availability of individual METFs estimated by Marquardt and Pedersen (2010) offers the possibility to examine how such differences in shape can affect the auditory filter characteristics at low CFs. In this study, individual differences in frequency

selectivity at low CFs are estimated by assuming individual METFs of the shape suggested by the work of Marquardt and Pedersen (2010) and Pedersen and Marquardt (2009). These results are compared with individual differences in frequency selectivity estimated psychophysically using the notched-noise method. Equivalent rectangular bandwidth (ERB) and asymmetry of the auditory filter have been considered in the analysis. Preliminary results from a parallel study concerning low-frequency psychophysical tuning curves and equal loudness contours are reproduced here to provide further evidence for the findings. Results are expected to contribute to the general understanding of individual differences in the perception of low-frequency noise.

2. Experimental design

2.1 Notched-noise masking experiment

Masked thresholds were obtained for tones positioned in spectral notches of notched noise maskers. The frequencies of the tones were f_s = 50, 63, 80, 100, and 125 Hz. The notch was positioned both symmetrically and asymmetrically around the tone. In the symmetric case, the normalized frequency separation, $\Delta f / f_s$, between the edge of the maskers and the tone were: 0, 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6. Two asymmetric notch conditions were included. In the 'upper+' case, the values of $\Delta f / f_s$ were 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6 for the lower band, while the corresponding values for the upper band were 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8. In the 'lower+' case, the values of $\Delta f / f_s$ were 0.1, 0.2, 0.3, 0.4, 0.5, and 0.6 for the upper band, while the corresponding values for the lower band were 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8. The spectral density level of the masker was N_0 = 50 dB for both bands when f_s = 80, 100, and 125 Hz. For f_s = 50, and 63 Hz, this level was N_0 = 50 dB for the upper band and N_0 = 62 dB for the lower band. As described by Jurado and Moore (in prep), applying emphasis to the lower band allows a better estimation of the lower auditory filter skirt.

Noise signals were 600 ms long, including 50 ms linear ramps at start and end. The tone signals were 500 ms long, with 25 ms linear ramps at start and end. Tone signals started 50 ms after the noise maskers and ended 50 ms earlier. A 3-down 1-up adaptive procedure, with a 3 alternative forced choice task, was used to estimate the 79% point in the psychometric function. The three noise bursts in a sequence were separated by 500 ms silence intervals. Two masked thresholds were obtained, and if their difference exceeded 3 dB a third measurement was done, and all (the 2 or the 3) were averaged.

The masking data was fitted in a power spectrum model of masking scenario in order to derive the auditory filter shapes (Fletcher 1940; Patterson et al. 1982) . A roex(p,r) auditory filter was used (Patterson et al. 1982). Both filter skirts were allowed to differ (the slope parameters are: p_u for the *upper* side and p_l for the *lower* side). For such a filter, the *r* parameter acts as a dynamic range limitator, and high values in the slope parameters indicate higher frequency selectivity and vice-versa. More details of the method are given in Jurado and Moore (in prep).

2.2 METF estimates from DPOAE

For details on the method for obtaining METF estimates from DPOAE measurements see Marquardt and Pedersen (2010) and Pedersen and Marquardt (2009). The frequencies for which the METF was estimated were 20, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 125, and 250 Hz.

The responses were interpolated in order to obtain high resolution versions which were used in the calculations. For purposes of calculation, extrapolation of the responses was done as well, estimating values down to 10 Hz, where the slope of each subject's METF was preserved.

2.3 Subjects

A total of 6 subjects participated in the notched-noise experiment. Four of these subjects were tested at all CFs, while two of them were tested only at 50 and 63 Hz. On the other hand, data from 5 subjects was obtained in the DPOAE measurements. Both groups, the 6 subjects from the notched-noise experiment and the 5 subjects from the DPOAE measurements, consist of a different set of subjects.

2.4 Deriving individual estimates of frequency selectivity

Individual estimates of frequency selectivity derived from the masking experiment described in section 2.1 were obtained considering the system as a whole (i.e. without correcting for an assumed METF). This was done because individual METF estimates of the subjects that participated in this experiment were not available. Thus, the fitted filter shape is used to account for the effects of the individual METFs in these cases. Results from these cases will be termed "direct psychophysical estimates" in the rest of the paper.

On the other hand, in order to estimate the effect of individualities in the shape of the METF on frequency selectivity, the 5 METFs obtained by Pedersen and Marquardt (2009) were used in combination with the auditory filter shapes described in the study of Jurado and Moore (in prep). These METFs were obtained on a different set of subjects; therefore results assume each group of subjects present representative variations in frequency selectivity and shape of the METF. The individual METFs and auditory filter shapes were cascaded, thus producing individual estimates of frequency selectivity considering the system as a whole (i.e. a set of "combined responses"), allowing comparison with the direct psychophysical estimates. The chosen underlying auditory filter parameters correspond to those derived from the mean masking data of Jurado and Moore (in prep) (which were obtained applying a 12 dB emphasis to the lower masking band at 50 and 63 Hz), after correction for their assumed "default" METF.

It is important to note that, in the region of effective influence of the helicotrema (i.e. below about 50 Hz), the cascading of an underlying filter and the METF (that includes the helicotrema shunt effect) has been done for calculation purposes. In this frequency region, the helicotrema shunt mechanism is thought to be dominant and to largely represent cochlear frequency selectivity; the underlying filters being merely representative. Indeed, the assumed

parameters do not contradict this notion, as the underlying filter shapes derived by Jurado and Moore (in prep) present increasingly shallow low frequency skirts (i.e. little frequency selectivity below the CF) as CFs approach 50 Hz.

The 5 individual METFs present different central locations and extent in the irregular-transition region separating the two different slope regions, therefore it seems a rather smooth-non-individual transition region in the METF, as was assumed by Jurado and Moore (in prep), was appropriate to analyze mean masking data. Consequently, the assumed auditory filter parameters are thought to represent mean underlying auditory filter parameters, and it is assumed that the individuality is obtained after filtering with the individual METFs.

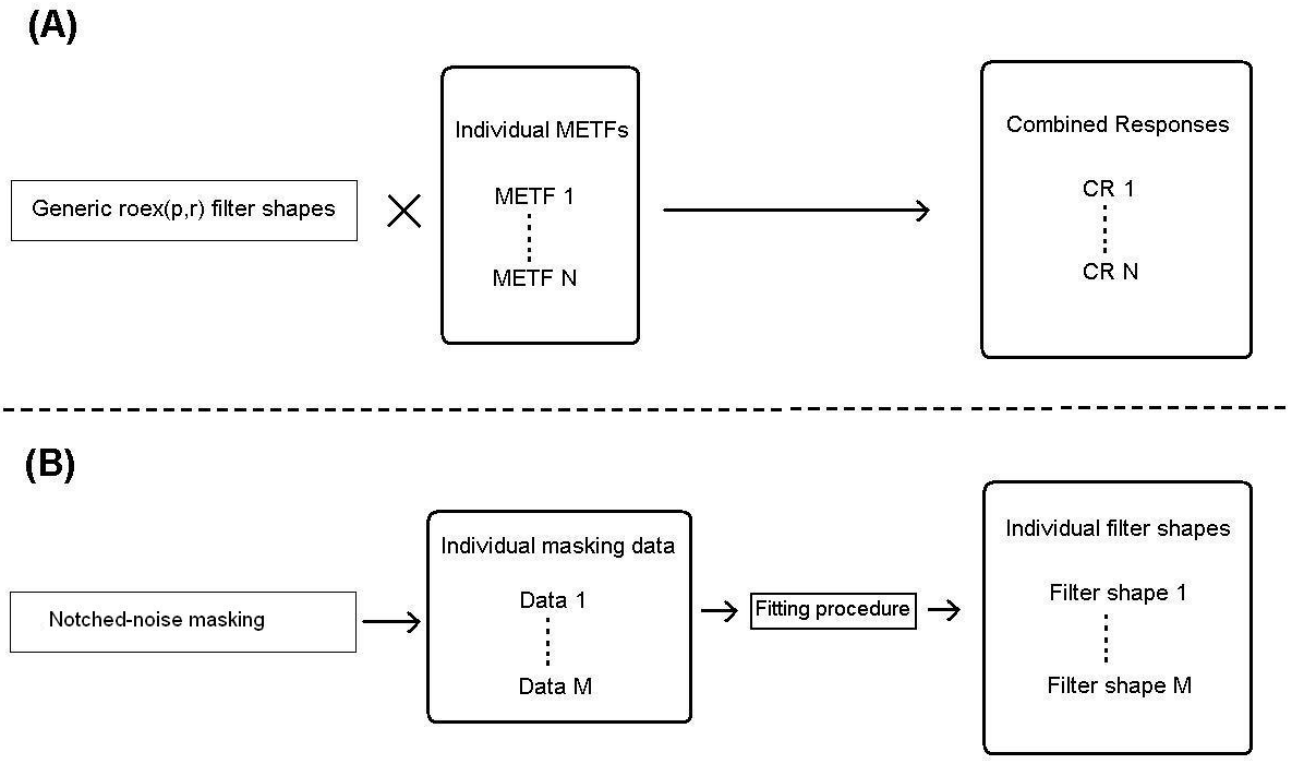

Figure 1: Block diagram of main procedures used to obtain (A): the “combined responses“ representing the individually filtered estimates; and (B): filter shapes representing the direct psychophysical estimates. N and M are the number of filter shapes obtained in each procedure (N = 5 and M = 6).

Each METF was cascaded with the assumed filter parameters at the CFs of 50, 63, 80, 100, and 125 Hz. After frequency domain multiplication, the combined responses were normalized to a 0 dB maximum gain after dividing by its maximum value, which was typically at the CF

(due to the filter attenuation at both sides off its centre). This was not the case in some instances for 63 and 50 Hz, and 80 Hz due to the steep gain increases (within the irregularity region) in the METFs around these very low CFs. The point of maximum gain (i.e. tip of the tuning curve) and the presence of inflection points in the combined responses will be used in the analysis to describe how tuning is affected at a particular CF. This is described in section 3.2. A set of individual auditory filter shapes, representing frequency selectivity of the system as a whole, was obtained after the cascading- and normalization procedure. These filter shape estimates will be called "individually filtered estimates" in the rest of the paper. Subsequently, numerical integration of the individually filtered estimates was performed in order to estimate the ERB at each CF. These ERBs and filter shapes were then compared with the ERBs and filter shapes obtained in the direct psychophysical estimates. Figure 1 summarizes the main procedures used to obtain the individually filtered estimates (A) and direct psychophysical estimates (B).

3. Results and discussion

3.1 Description of the individual METFs

Two examples of individual METFs obtained by Marquardt and Pedersen (2010) are reproduced here in figure 2. For ease in the interpretation of the results, the METFs were described according to 3 criteria as indicated in the figure. These are: 1) The frequency corresponding to the maximum amplitude in the elevation at the low-frequency side of the depression; 2) The frequency of the dip (minimum amplitude) in the depression of the METF ; and 3) The upper cut-off frequency where the steeper inner slope (i.e. slope within the irregularity) of the METF ends. The largest differences occurred in criteria 3. Some of the METFs presented more clear transitions out from the irregular region than others, therefore criteria 3 was as well more difficult to establish. The METFs shown in figure 1 (labelled numbers 2 and 4) are two representative examples, which appear to differ particularly in the later criteria. Table 1 shows the frequency values corresponding to criteria 1-3 for each METF. The extent of the irregularity region, described here as the difference between the frequencies corresponding to criteria 3 and 1, is 35,35,45,70, and 80 Hz for METFs # 1,2,3,4, and 5, respectively.

Table 1: Frequencies (Hz) corresponding to criteria 1,2, and 3 for each of the 5 METFs.

	Criteria (Hz)		
METF #	**1**	**2**	**3**
1	45	57.5	80
2	40	50	75
3	35	40	80
4	55	65	125
5	45	55	125

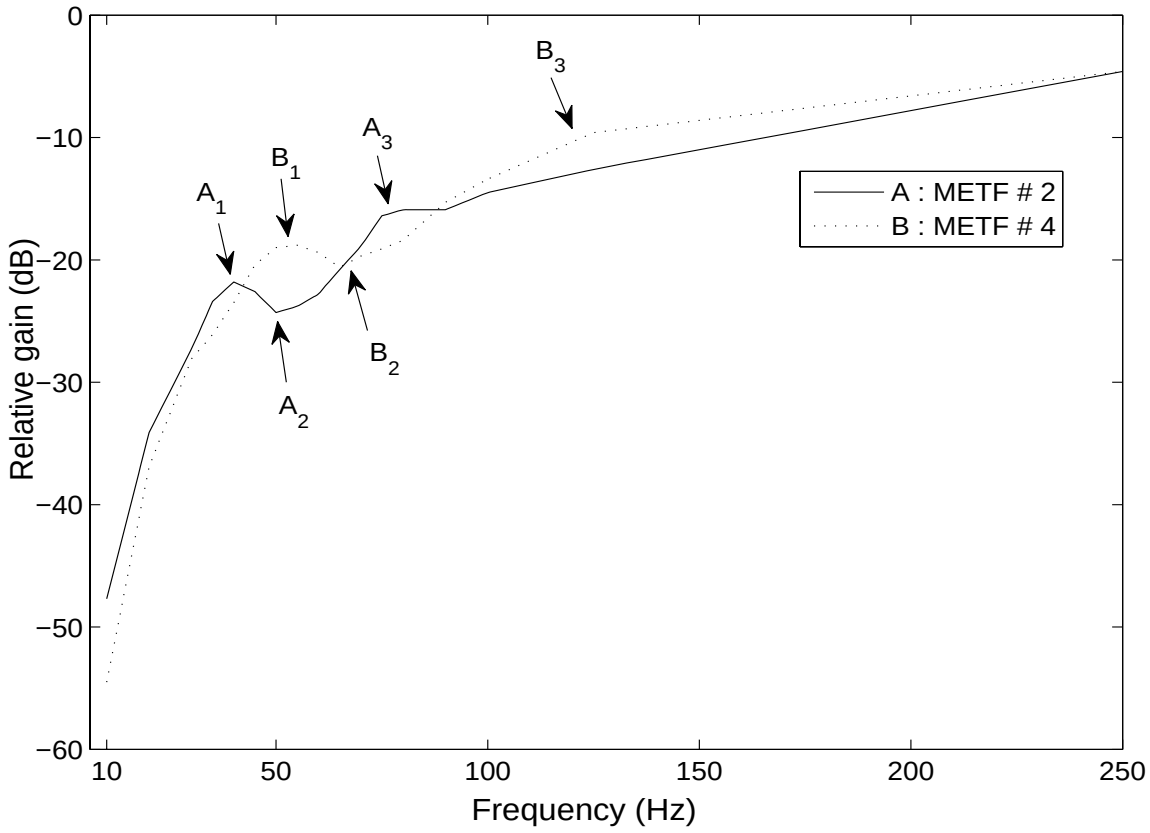

Figure 2: Example individual METFs (METFs # 2 (A) and 4 (B), solid and dotted lines, respectively). For comparison purposes the gain is set relative to an arbitrary reference of 0 dB at 1000 Hz, and assuming in each case the gain is -4.6 dB at 250 Hz relative to 1000 Hz, as it is the case for the METF used in the ANSI standard for loudness calculation (ANSI 2007). The frequencies corresponding to criteria 1 to 3 are denoted by A_1, A_2, and A_3 and B_1 , B_2, and B_3, and indicated by the arrows. These are respectively 40, 50, and 80 Hz for cases (A) and 55, 65, and 125 Hz for cases (B).

3.2 Individual auditory filter shapes representing frequency selectivity of the system as a whole

Figure 3 shows the auditory filter shapes obtained for the individually filtered (left) and direct psychophysical estimates (right). As the figure shows, at all CFs the overall-filter shape characteristics of both estimates present similar features. This is true even though the assumed underlying auditory filter shapes (dotted lines, left side) are quite different from the “combined responses”, especially at the lowest CFs. The degree of individual differences seems as well comparable taking into account both estimates come from different groups of subjects.

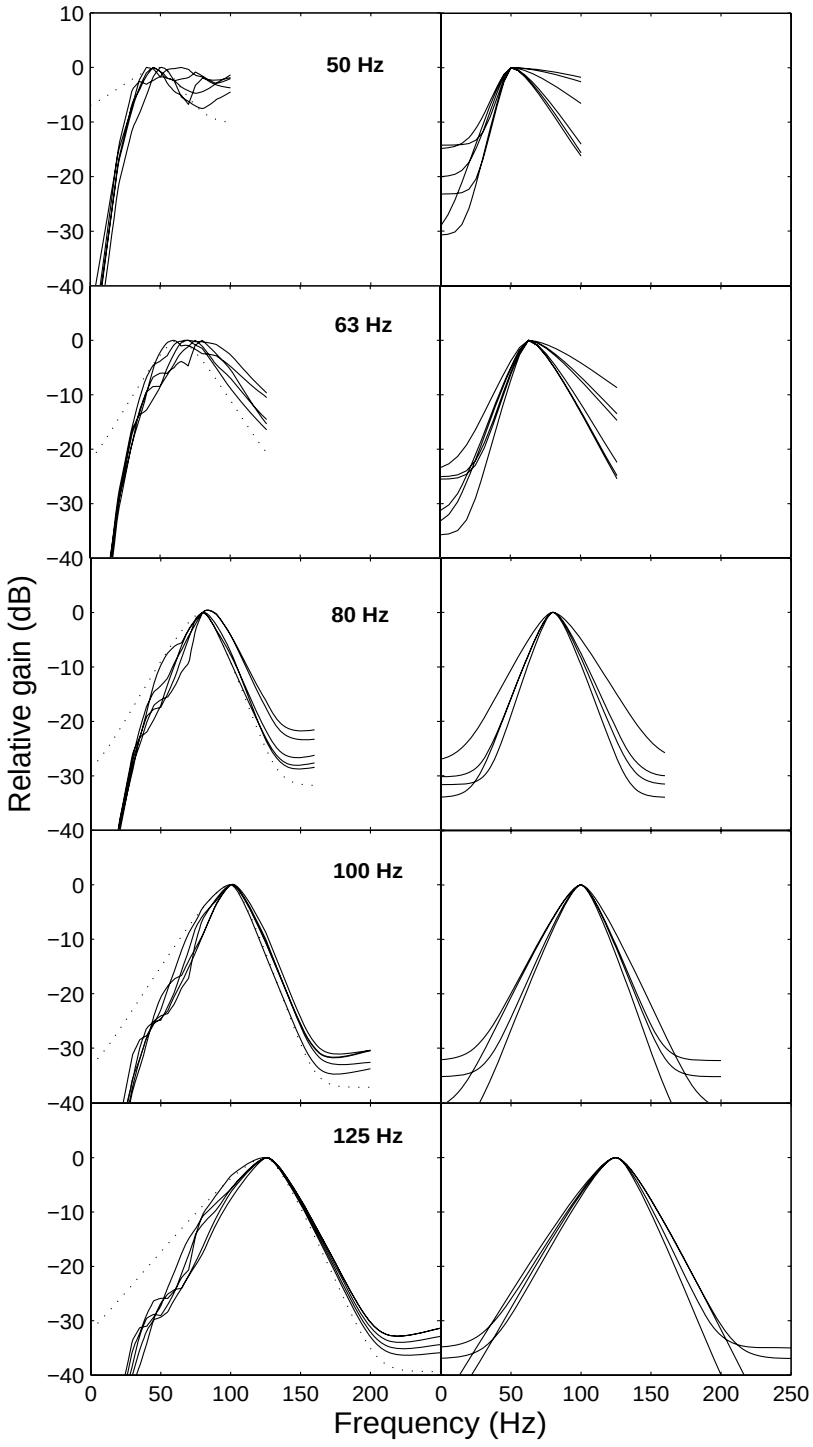

Figure 3: Individual auditory filter shapes obtained for the individually filtered estimates (solid lines left) and direct psychophysical estimates (solid lines right). The dotted lines on the left side are the underlying auditory filter shapes assumed at each CF.

While the hard dynamic range limitation of the roex(p,r) filter affects both sides of the filter skirts, the large attenuation assumed at the low-frequency side in the METFs does not have this imitation. This effect explains differences in the dynamic range at the low frequency side of the filter skirts, evident for the lowest CFs. An apparent worsening of frequency selectivity can be appreciated in both groups of filter shapes at 63 and 50 Hz, respect to the higher CFs. For 125, 100 and 80 Hz, the irregular region in the METFs affects only marginally the lower skirt of the individually-filtered filter shapes and in both estimates the filters appear to be more symmetrical. However, for 63 and 50 Hz, both estimates show clearly steeper lower filter skirts than upper skirts. This effect was previously reported by Jurado and Moore (in prep) from analysis on the mean masking data and was attributed to the increasing steepness of the METF below about 50 Hz, due to the helicotrema shunt action. For the direct psychophysical estimates, the standard deviations (std) in auditory filter asymmetry (p_u/p_l) were calculated, and these were similar across frequency, though perhaps somewhat larger for 50 Hz (std values of 0.16, 0.14, 0.10, 0.09, and 0.13 for 50, 63, 80, 100, and 125 Hz, respectively).

For 63 Hz and especially for 50 Hz, frequency selectivity above CF is quite poor, with relatively large individual differences, present in both, individually filtered and psychophysical estimates. In some cases the upper skirt is almost completely shallow (almost no selectivity) and in others it is relatively much steeper. Some very low values of p_u, such as values close to 1, could be observed in the direct psychophysical estimates. Inspection of the individual shape of the METFs indicated that for the individually filtered estimates, the large flattening off in the upper filter skirts was a consequence of the shape of the irregularity region, which presents steep changes in slope and a change in slope direction (i.e. an inflection point), which even created increases in gain above CF. It is noteworthy that the roex(p,r) fits to the psychophysical data could estimate as well these almost completely shallow slopes, although the filter by default is limited to skirts of exponential shape, and no individual particularities can be modelled. The observed individual differences indicate that for some subjects, low frequency signals (with components around e.g. 50 Hz) will create much larger masking effects, and presumably a larger internal excitation, than for others.

A more detailed description of how the METF influences the shape of the combined response (CR) is given in figure 4. Here the CRs obtained for METFs # 2 and # 4 are plotted for the CFs of 50, 63, and 80 Hz. As can be appreciated, the somewhat steep positive slope region within the irregularity is responsible for the large flattening off observed in the upper filter skirts. If a filter is assumed to be centred within this region, its negative slope above CF cannot effectively tune down and counteract this effect, thus the resulting upper filter skirt will be very shallow (e.g. see upper figure for CF = 50 Hz) or even present an inflection point (e.g. see lower figure for CF = 50 Hz). Furthermore, the manner in which different CFs are affected appears to depend on their relative locations respect to the irregular region, which by itself presents rather individual characteristics (e.g. criteria 1-3 frequencies).

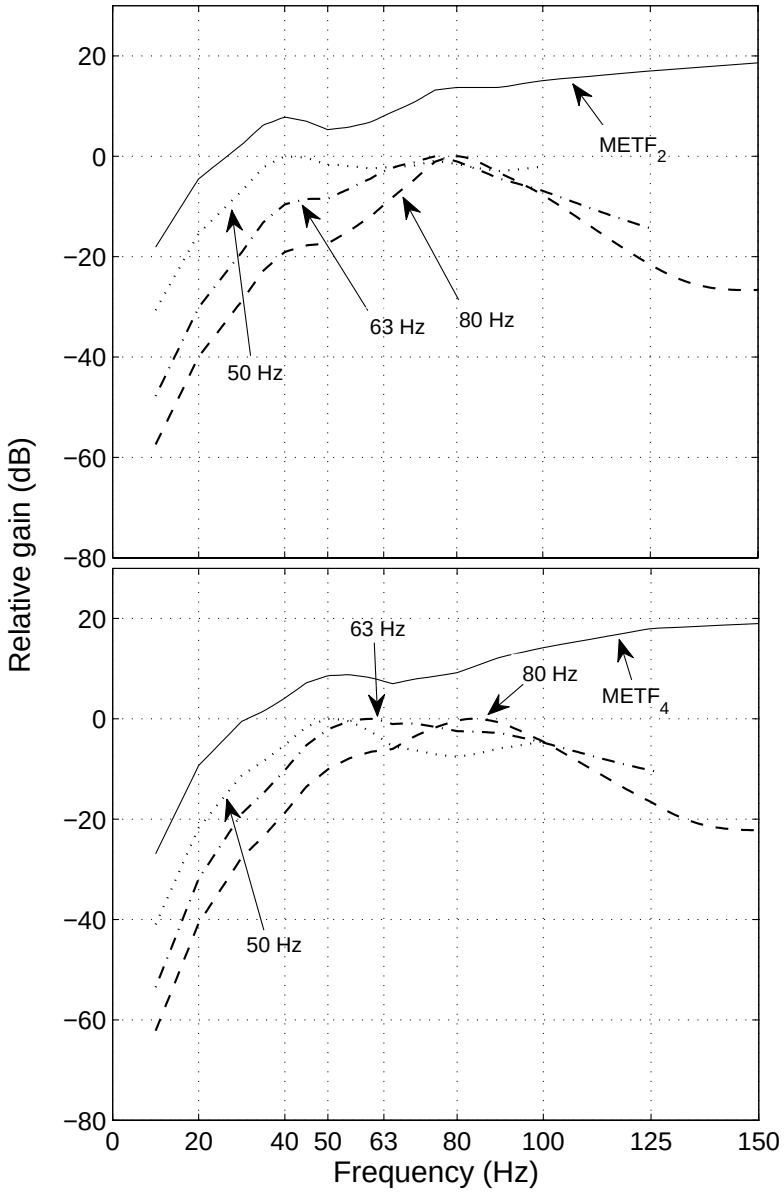

Figure 4: Combined responses # 2 and 4 (upper and lower panels, respectively) for the CFs of 50, 63, and 80 Hz (dotted, dashed-dotted and dashed lines, respectively and indicated in the figure by the arrows). The responses have been normalized by their maximum gains. The corresponding METFs (# 2 and 4) are plotted in solid lines at arbitrary reference for ease in the comparison.

For example, for the CF of 63 Hz, the maximum gain is located well above CF (~75 Hz) for the CR # 2 and below and above this point the filter appears to tune (i.e. the filter skirts produce attenuation at both sides respect to the tip). On the other hand, for CR # 4 the 63 Hz

filter has its maximum gain near the CF (~60 Hz), but now its upper skirt is much shallower because it falls fully in the steeper region of the irregularity. A similar reasoning explains that the maximum gain of the filter for the CF of 50 Hz is at a higher frequency for the case of CR # 4 (~50 Hz) than for CR # 2 (~40 Hz), since METF # 4 presents a higher criteria 2 frequency (i.e. frequency of the elevation at the low-frequency side of the depression). Similarly, due to the higher criteria 3 frequency of METF # 4 (i.e. its irregular region appears to extend up to higher frequencies), even the CF of 80 Hz was affected for CR # 4, since the maximum gain of the filter is located at about 84 Hz. This is not the case for METF # 2, where the maximum gain is at 80 Hz.

Table 2 shows the frequencies where the maximum gain of the CR is located for the 5 cases (in respective numbering as for the METFs). The frequencies of the inflection points (i.e. change in slope direction) are also shown in parenthesis where applicable.

Table 2: Frequencies of maximum gain for the CRs (Hz). If the CR contains an inflection point, the frequency if the inflection point is shown in parenthesis.

	CF (Hz)				
CR #	**50**	**63**	**80**	**100**	**125**
1	45 (70)	80	81	100	125.5
2	40(60)	75	80	100.5	126
3	65	70	80	100.5	126
4	50.5(80)	60	83.5	101.5	125.5
5	45 (75)	69	83.5	101	125.5

The maximum gains for the CF of 50 Hz was in most cases close to the frequency corresponding to criteria 1. However, nearly all the CRs for the CF of 50 Hz present inflection points above the CF, which in all cases fall within the irregularity region (i.e. between the frequencies of the METF corresponding to criteria 1 and 3). The maximum gains for the CF of 63 Hz are close to the frequencies of criteria 3 of METFs # 1, 2, and 3 for CRs # 1, 2, and 3, and within the irregularity region of METFs # 4, and 5 for CRs # 4 and 5 (i.e. in most cases they are above CF). For the CFs of 80 Hz and above, the maximum gains are within 1.5 Hz from the CF, excepting for CRs # 4 and 5, at 80 Hz, which present a somewhat higher deviation (3.5 Hz above). Since the irregularity region of METFs # 3 and 4 seems to extend to somewhat higher frequencies than for the other cases (higher criteria 3 frequency), it makes sense that its steepness could still influence tuning of the filter up to the CF of 80 Hz. In overall, this suggests that for CFs that fall within the irregularity region, a “tip shift”, typically upwards in frequency, will occurr.

3.3 Equivalent rectangular bandwidths

The ERB values obtained for the direct psychophysical and individually filtered estimates are shown in figure 5. Interestingly, the ERB values and spread correspond relatively well between the 2 groups of estimates. For the direct psychophysical estimates, standard deviations of 8.9, 7.5, 5.2, 2.7 and 1.9 Hz were respectively obtained for 50,63,80,100, and 125 Hz. In the case of the individually filtered estimates, these were respectively 8.4, 9.8, 6.1,

2.7, and 4 Hz. In the direct psychophysical estimates, larger ERB values were all correlated with lower values of p_u, as expected from the ERB calculation.

It appears that the shape of the individual METFs can largely account for the observed individual differences in frequency selectivity. The individually filtered estimates produce ERBs that follow the general trend of increasing below about 80 Hz as with the direct psychophysical estimates (this was reported previously by Jurado and Moore (in prep) from analysis on the mean masking data). Individual differences in this trend are also accounted, such as cases where the increase happens down from 80 Hz, or cases where an increase or decrease in ERB occurs at 50 Hz relative to 63 Hz. Interestingly, cases where a relative increase in the ERB occurred already at 80 Hz for the individually filtered estimates correspond to CR # 4 and 5. As seen in the previous section, both corresponding METFs have irregularity regions which appear to extend above 80 Hz.

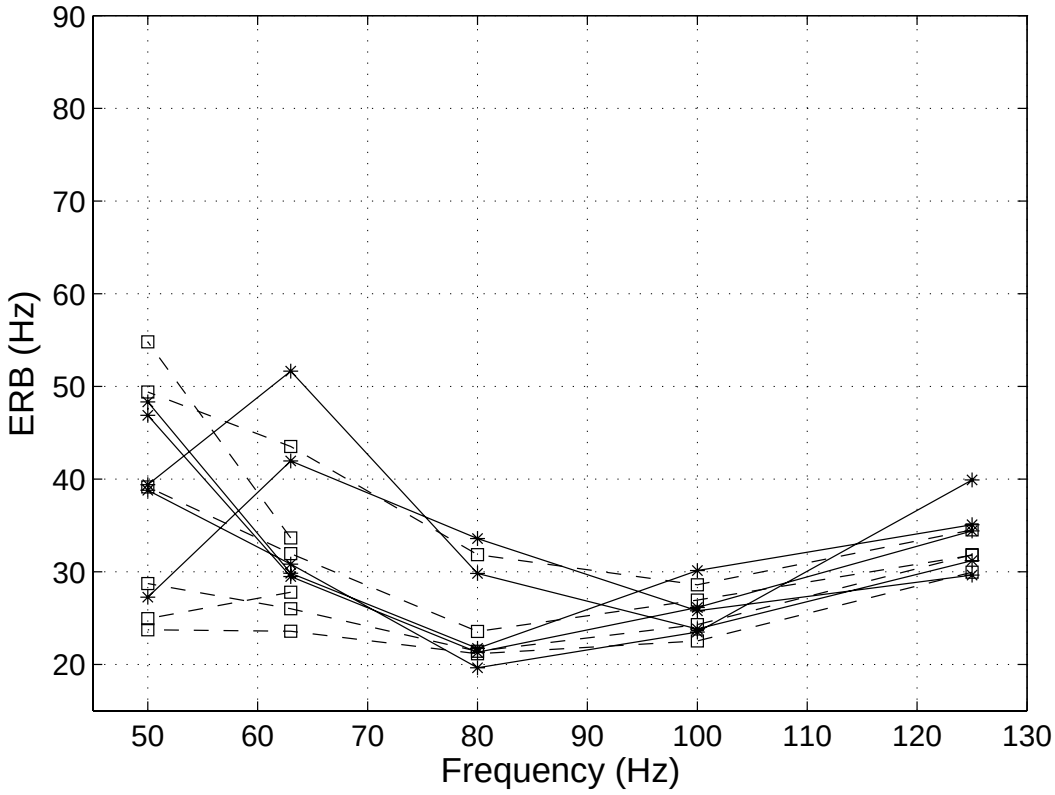

Figure 5: ERB values corresponding to the direct psychophysical estimates (squares and dashed lines) and to the individually filtered estimates (asterisks and solid lines).

From the description provided in section 3.2, it seems now evident that the irregularity region, due to its steep positive slope region, affects tuning of the auditory filter for CFs that fall within

this region. This appears to explain not only the filter shape characteristics observed in the direct-psychophysical estimates (described in section 3.2), but as well, the observed general trends in the ERB.

3.4 Comparisons with preliminary results from a parallel study treating psychophysical tuning curves and equal loudness contours in the low frequency region

Preliminary results from a parallel study are reproduced here to facilitate comparisons of the main findings described in the previous sections. More details of the methodological approaches used in these experiments are given in Jurado et al. (2010).

Figure 6 shows example equal loudness contours obtained for 3 subjects using comparison tone frequencies of : 20, 30, 35, 40, 45, 55, 60, 65, 70, 75, 80, 90, 100, 125, and 160 Hz. A simple up down adaptive procedure has been used, with interleaved 2 track playback. The reference tone was a 50 Hz sinusoid at a fixed level of 40 dB SL. As can be seen, an irregular region, as observed in the METFs described earlier, is evident in these 3 cases. For one pilot-test subject, a 90 Hz comparison tone was used as well, and the same irregular region was still observed. The general trends are similar to those observed in the METFs described in section 3.1. As well, the detailed structure of the transition region appears to be rather individual, suggesting substantial individual differences in loudness perception within this region. These preliminary results also indicate that the shape of the METFs, measured objectively by the DPOAE method, appears to have a relevant perceptual influence. It should be noted that in the study of Pedersen and Marquardt (2009), where METFs and ELCs (for a 100 Hz, 20 dB SL comparison tone) were estimated for the same subjects, in only 2 cases a clear correspondence in the resonance features between the METFs and ELCs were found. Nevertheless, the following comments on these preliminary results will assume the inverse of the ELC is a rough estimate of the shape of the subjects' METF.

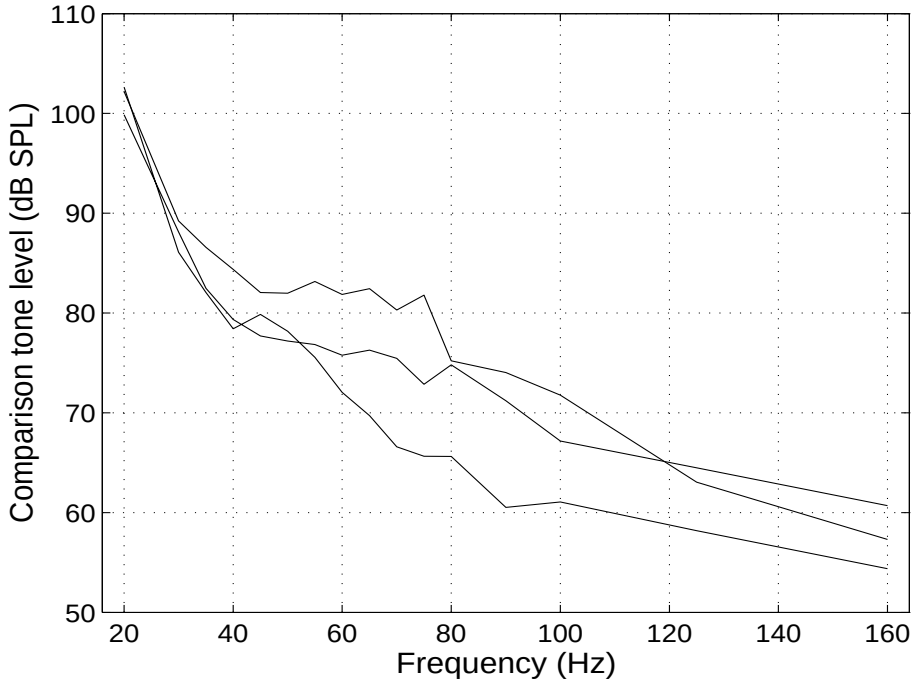

Figure 6: Example equal-loudness-level contours obtained at closely spaced frequencies for 3 subjects in the low-frequency region. At 50 Hz, the 40 dB SL is plotted.

Figure 7 shows example psychophysical tuning curves (PTCs) from 1 subject. The spectrum density level of a noise band (of bandwidth equal to 0.5×CF) required to mask a 15 dB SL sinusoid of frequency equal to the CF has been found. The CFs tested were 31.5, 40, 50, 63 and 80 Hz. A 0.2×CF resolution in the PTC has been used around the CF, and the step has been increased to include larger deviations from the CF. The curves have been inverted and normalized by their maximum values for comparison purposes with figure 4 (see section 3.2).

The pattern of results shown in figure 7 appears to be broadly in agreement with those shown in figure 4 (see section 3.2). The upper panel shows the same CFs as those evaluated in the individually filtered estimates. The upper skirt of the filter for CFs of 50 and 63 Hz appears to be affected, approximately within the irregularity region of the assumed METF, by the larger steepness of the irregularity, which seems to extend at least up to 100 Hz. Roughly above this region a higher tuning is achieved on the upper side of the PTC.

The lower panel in figure 7 shows results for the even lower CFs of 31.5 and 40 Hz. Here the influence of the shape of the METF on the PTC appears to be evident. The filters clearly do not reach their maximum gains, i.e. do not have their tips, at the CF. As described previously in section 3.2, at the lower side, the increasing steepness in the METF largely determines frequency selectivity. As well, an inflection point within the frequency region of the irregularity, previously described in figure 3 (observed at 50 Hz), can be observed for both of these CFs.

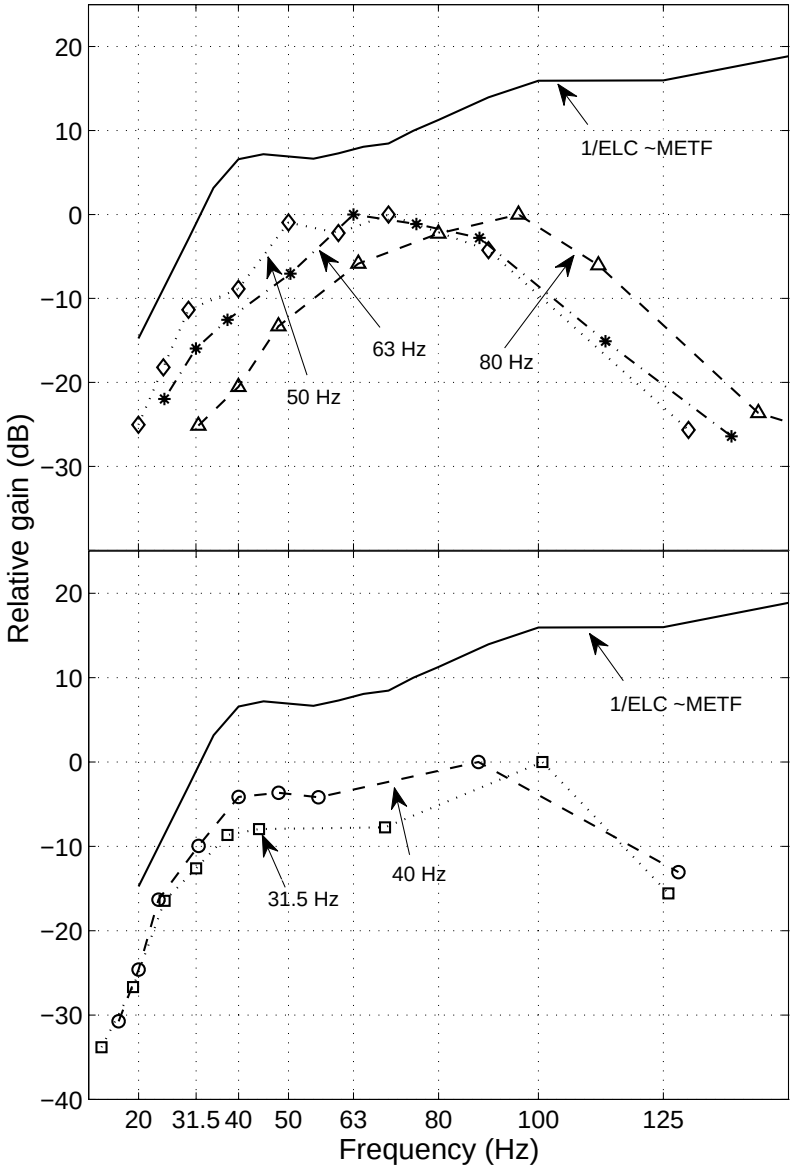

Figure 7: Example PTCs obtained for an individual subject. The upper panel shows the PTCs for the CFs of 50, 63, and 80 Hz (dotted line-diamonds, dashed-dotted line-asterisks, and dashed line-triangles, respectively). PTCs for the CFs of 31.5 and 40 Hz are plotted in the lower panel (dotted line-squares and dashed line-circles, respectively). Each point in the PTCs corresponds to the centre frequency of the band of noise. The ELC of the subject has been inverted and plotted (in solid line) at an arbitrary reference, for use as a rough indicator of the shape of the subjects' METF.

4. Further development

A more complete description, considering individual data, of the factors affecting frequency selectivity at these very low-CFs is thought to require detailed individual measures of

frequency selectivity together with estimations of each of the subject's METF. This is part of a current study where experiments are being carried out for this purpose (see Jurado et al. 2010). Psychophysical tuning curves, providing direct measures of tuning of the system as a whole, are considered at the 1/3 octave CFs of 31.5, 40, 50, 63 and 80 Hz. This method gives direct estimates of the tuning properties of the system, and will easily determine if no tuning occurs at a specific CF (i.e. see (Kluk and Moore 2005)). As well, an equal-loudness-level contour at closely spaced frequencies in the frequency region below 100 Hz, is measured for every subject in order to provide an estimate of the subject's METF, useful for analysis purposes. Objective measures of the subject's METF by using DPOAE techniques are being considered as well.

5. Conclusions

The evidence gathered in this work suggests that frequency selectivity is largely affected by the shape of the METF for CFs in the range below about 80 Hz. In overall, observed Individual differences in ERB and asymmetry of the auditory filter were larger for CFs that would typically fall within the irregular-transition region from approx 6 dB/oct to 12 dB/oct zones in the METF (such as 50 and 63 Hz) than for higher CFs. This transition region, and the steeper region below, is thought to be largely defined by the helicotrema impedance (Marquardt and Hensel 2008; Marquardt and Pedersen 2010). Thus, differences in helicotrema size and shape are possible reasons for the observed individual differences in frequency selectivity. Consequently, in the range of frequencies below about 100 Hz, the helicotrema appears to become an important factor that determines individual differences in masking and excitation patterns created by low-frequency noise.

The calculations using real METFs indicated that the large flattening off in the upper filter skirt, observed at 63 and 50 Hz, was due to the steep region within the irregularity in the METF. It could be predicted as well that this region may produce a shift in the location of maximum gain (i.e. in the tip) of auditory filters centred within it. The predicted-flattening off of the upper filter skirts agreed with that observed in notched-noise-derived filter shapes, and with the trends given in the example low-frequency PTCs. This was true, even though a fixed set of parameters representing a "mean cochlear roex filter" -obtained from the results (after METF correction) of Jurado and Moore (in prep)- was somewhat arbitrarily assumed at each CF. The flattening off in the upper filter skirt produced an increase in ERB with decreasing CF, starting from below about 80 Hz. Thus, it appears that the helicotrema has evolved to favour a reduction in low-frequency sensitivity (and therefore protection form large pressures and reduction of internal physiological noise) at the expense of a worsening of tuning.

References

ANSI (2007). ANSI S3.4-2007. Procedure for the computation of loudness of steady sounds. New York, American National Standards Institute.

Fletcher, H. (1940). "Auditory patterns." Reviews of Modern Physics **12**: 47-65.

Glasberg, B. R. and B. C. J. Moore (1990). "Derivation of auditory filter shapes from notched-noise data." Hearing Research **47**: 103-138.

ISO 226 (2003). Acoustics - normal equal-loudness contours. Geneva, International Organization for Standardization.

Jurado, C. A. and B. C. Moore (in prep). "Factors affecting frequency selectivity for frequencies below 100 Hz: comparisons with mid-frequencies." Journal of the Acoustical Society of America.

Jurado, C. A., B. C. J. Moore and C. S. Pedersen (2010). Psychophysical tuning curves for very low centre frequencies. Proceedings of 20th International Congress on Acoustics, ICA 2010 Sydney, Australia.

Kluk, K. and B. C. J. Moore (2005). "Factors affecting psychophysical tuning curves for hearing-impaired subjects." Hearing Research **200**: 115-131.

Leventhall, H. G. (2004). "Low frequency noise and annoyance." Noise and Health **6**: 59-72.

Marquardt, T. and J. Hensel (2008). A lumped-element model of the apical cochlea at low frequencies. Proceedings of the10th International Workshop on the Mechanics of Hearing, Keele, U.K.

Marquardt, T., J. Hensel, D. Mrowinski and G. Scholz (2007). "Low-frequency characteristics of human and guinea pig cochleae." Journal of the Acoustical Society of America **121**(6): 3628-3638.

Marquardt, T. and C. S. Pedersen (2010). The infuence of the helicotrema on low-frequency hearing. The neurophysiological bases of auditory perception. E. A. Lopez-Poveda, A. R. Palmer and R. Meddis. New York, Springer.

Møller, H. and C. S. Pedersen (2004). "Hearing at low and infrasonic frequencies." Noise & Health **6**(23): 37-57.

Moore, B. C. J., B. R. Glasberg and T. Baer (1997). "A model for the prediction of thresholds, loudness and partial loudness." Journal of the Audio Engineering Society **45**: 224-240.

Moore, B. C. J., R. W. Peters and B. R. Glasberg (1990). "Auditory filter shapes at low center frequencies." Journal of the Acoustical Society of America **88**: 132-140.

Patterson, R. D., I. Nimmo-Smith, D. L. Weber and R. Milroy (1982). "The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold." Journal of the Acoustical Society of America **72**: 1788-1803.

Pedersen, C. S. and T. Marquardt (2009). Individual differences in low-frequency noise perception. INTER NOISE. Otawa, Canada.

Rosen, S. and D. Stock (1992). "Auditory filter bandwidths as a function of level at low frequencies (125 Hz-1 kHz)." Journal of the Acoustical Society of America **92**: 773-781.

Unoki, M., I. Kazuhito, Y. Ishimoto and C. T. Tan (2006). "Estimate of auditory filter shape using notched-noise masking for various signal frequencies." Acoustical Science and Technology **27**: 1-11.

Zwicker, E., G. Flottorp and S. S. Stevens (1957). "Critical bandwidth in loudness summation." Journal of the Acoustical Society of America **29**: 548-557.

The effect of the helicotrema on low-frequency hearing: II. Equal-loudness contours

Carlos A. Jurado
Section of Acoustics, Department of Electronic Systems, Aalborg University, Fredrik Bajers Vej 7-B5, 9220 Aalborg Ø, Denmark

Torsten Marquardt
UCL Ear Institute, 332 Grays Inn Rd London, WC1X 8EE, UK

Abstract

Equal-loudness-level contours (ELCs) were measured for 14 subjects at closely spaced frequencies between 20 and 160 Hz. These were compared with distortion-product-isomodulation curves (DPIMC) determined for both ears of each subject. The latter functions estimate the shape of the forward middle-ear transfer function, which includes the effects of the helicotrema shunt mechanism. A close match between the shape of the ELCs and the DPIMCs was generally found, particularly within an irregularly shaped region between about 40 to 100 Hz. For frequencies below 40 Hz, applying a 6 dB/oct increase in slope to the DPIMCs accounted generally well for the perceptual ELC data. Results suggest that perceived loudness for sinusoidal signals is largely determined by cochlear anatomy. Besides, individual differences in perpeption may easily arise due to pronounced individual differences observed in the shape of the DPIMCs, particularly for frequencies between about 40 to 100 Hz. Results are discussed in terms of their relation to standardized equal loudness contours.

I. INTRODUCTION

Much of the sound in our every day environment contains considerable energy in the low-frequency range. Such sounds can easily propagate and produce annoyance (Leventhall, 2004). An adequate characterization of our perception of low-frequency sounds is key to understand and assess problems produced by low-frequency noise. For low-frequency sinusoidal signals, audibility and loudness can be predicted by the use of standardized threshold and equal-loudness contours (see e.g. ISO 226, 2003). However, some evidence exists that within the region below 100 Hz, these curves may not be fully representative of an individual's perception (Marquardt and Pedersen, 2010; Marquardt *et al.*, 2007; Frost, 1987; Møller and Pedersen, 2004). Marquardt *et al.* (2007), estimated the shape the forward middle-ear-transfer function (fMETF) non-invasively by means of distortion-product-isomodulation curves (DPIMCs). The shape of the latter curves was assumed to correspond to that of the inverse of the fMETF, which inlcudes the effects of the helicotrema shunt mechanism. Their results suggested that there is an irregular region in the fMETF between about 40 to 100 Hz, that separates a shallower slope region above that (about 6 dB/otc) and a steeper region below (about 12 dB/oct). This region had the shape of a small resonance feature, and it was anticipated that it may have an influence in the perception of low-frequency sound.

Subsequently, Marquardt and Pedersen (2010) measured both equal-loudness contours (ELCs) and DPIMCs for a group of 5 subjects. Their DPIMCs showed substantial individual differences in shape

and extent of the irregular transition region. However, for all individuals, the DPIMCs were similar between the two ears. Only in two out of five cases they found a close match between the shape of the irregular region in the DPIMCs and the ELCs. In spite of this fact, the evidence gathered by them suggested that there may be in fact pronounced individual differences in cochlear anatomy, especially between about 40 to 100 Hz, which may lead to marked individual differences in perception. The latter was anticipated by Jurado et al. (2010b), who estimated the effects of the shape of individual estimates of the fMETF on auditory tuning. From this perspective, a further investigation has been considered necessary.

In the present work, equal-loudness contours have been obtained for 14 subjects for closely spaced frequencies between 20 and 160 Hz. These have been compared to DPIMCs measured at both ears of each subject. The latter are treated in detail in a companion manuscript in preparation (Marquardt and Jurado, 2010) together with fits from a model of the apical cochlea. Comparisons were done to investigate the similarity between the ELCs and DPIMCs and the possible dependence of loudness perception on cochlear anatomy.

II. METHOD

A. Equal-loudness-level contours

An equal-loudness-level contour was obtained for each subject in the frequency range from 20 to 160 Hz. The reference signal was a sinusoid with a frequency of 50 Hz and a fixed level of 40 dB sensation level (SL). Comparison sinusoids were used with frequencies of 20, 30, 35, 40, 45, 55, 60, 65, 70, 75, 80, 90, 100, 125, and 160 Hz. All signals were 1.2 sec long including 0.2 sec linear ramps at start and end.

1. Measurement of absolute thresholds

The absolute threshold of each subject was determined for a 50 Hz sinusoid, to set it as a reference level for the subsequent 40 dB SL presentation in the loudness experiment. This measurement was done just before the loudness experiment started. The duration and ramp characteristics of the signal were the same as for the ELCs, as described above. A 3-alternative forced-choice task with a 3-down 1-up adaptive procedure was used to estimate the 79% point on the psychometric function (Levitt, 1971). The silent interval between the three stimuli in a trial was 400 ms. Feedback was provided after each response. The procedure started with a simple 1-down 1-up rule for the first four presentations in order to rapidly approach the region of absolute threshold. The starting level of the signal was always 15 dB above the ISO-226 (2003) absolute threshold levels to ensure it was clearly audible at the beginning. The stepsize started at 8 dB, was decreased to 4 dB after 2 reversals, and was decreased to 2 dB after 2 further reversals, where it remained.

A total of 12 reversals was obtained and the absolute threshold was estimated as the average of the levels at the last 8. Two threshold estimates were obtained. If the difference between them was more than 3 dB, a third estimate was obtained. The two closest values were averaged to estimate the absolute threshold.

2. Loudness balance procedure

To estimate the 50 % point in the psychometric function corresponding to equal loudness, a 1-up 1-down, adaptive loudness balance procedure was used. The 50 Hz reference (fixed-level) tone and comparison tones were presented in random order in a 2-alternative task. The silent

interval between presentations was 400 ms. Two interleaved tracks were used. The procedure randomly selected one of the two tracks when choosing the level of the next presentation. For a given comparison tone, the level of one track started 10 dB below the 40 phon standardized ISO equal loudness level (ISO 226, 2003), and the other track started 10 dB above it (the ISO curve was interpolated to obtain values at all frequencies). This was done except when the frequency of the comparison tone was 20 Hz, in which case the tracks started at ±5 dB from the 40 phon level, avoiding initial levels that were too high. These starting levels were found to be adequate after pilot testing; typically the listener's point of subjective equality lay between these extremes. The stepsize started at 8 dB, was decreased to 4 dB after 2 reversals, and was decreased to 2 dB after 2 further reversals, where it remained. For each track, 10 reversals were obtained and the point of subjective equality (PSE) was obtained from the average level at the last 6. For a given run (i.e. single measurement), one PSE was determined, corresponding to the average PSE of the two tracks. Two runs were performed, and the PSE was estimated from the average PSE for the two runs. If the two PSE estimates differed by more than 3 dB, a third run was performed and all three PSE estimates were averaged. The choice of the next comparison signal to be tested was randomized. Breaks took place regularly after about 5 to 8 minutes of measurement to maintain concentration. The ELC for each subject was obtained in a single session during one day, lasting about 2 ½ hours in total including breaks.

B. Apparatus

A cabin, especially designed for playback of low-frequency signals under pressure field conditions was used. The cabin's inner dimensions are 0.8 × 1.4 × 0.9 m (W x H x D ~1 m^3). It has four Seas 33 F-WKA 13-inch loudspeakers in each side wall, positioned behind covering panels. The listener sat facing the door, with the ears at 90° relative to the side walls. The cabin provides an effective pressure field in the overall volume up to about 61 Hz.

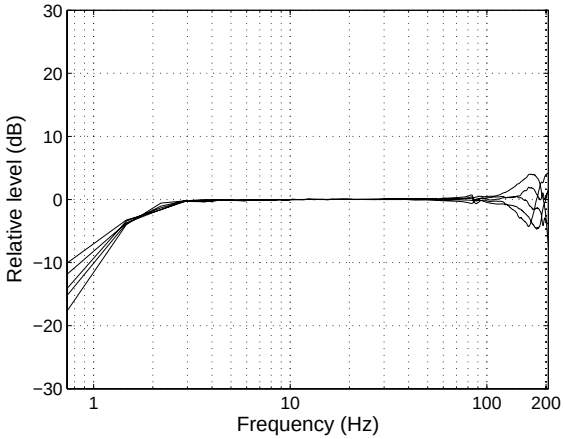

Figure 1. Differences between the magnitude responses at each of the five listening-positions and the average magnitude response of these positions in the listening cabin.

However, deviations in space are small up to higher frequencies (below ±3 dB up to about 150 Hz) within a defined listening space. A listener's position was defined by 5 points covering a region where it was expected that the listener's ears and centre of the head would be located. Equalization was done by filtering all stimuli with a digital filter representing the inverse room magnitude response, obtained after averaging the magnitude responses measured at the 5 positions. When performing the task, all listeners were instructed to find a comfortable position, without leaning to the sides, and to avoid moving much within blocks of measurements. Figure 1 shows the differences between the assumed magnitude response (0 dB reference), and the magnitude responses at the 5 positions, as a function of frequency. One-third octave

background noise sound pressure levels were measured in the cabin. Equivalent levels were obtained in 4 min 10 sec intervals and for worst-case scenarios, when activity was expected nearby in the building. The levels were measured with a low-noise microphone (B&K 4179, noise floor −5.5 dBA) using a 01 dB Harmonie system frequency analyzer. A worst case scenario example is shown in figure 2. As illustrated, all third-octave noise levels in the low-frequency range are well below the hearing threshold. The overall background noise level in the cabin was 21 dB SPL (~0 dBA) over the frequency range 5-20,000 Hz.

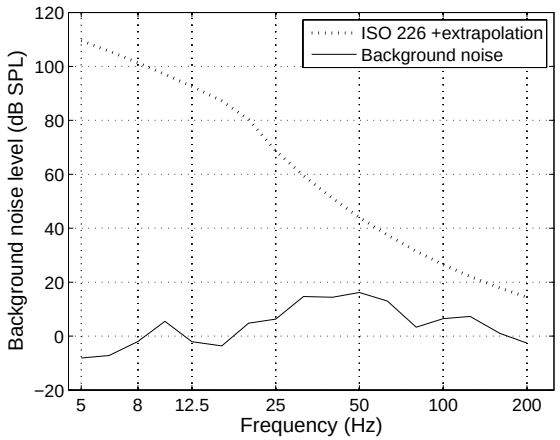

Figure 2. One-third octave background noise SPLs in the cabin (solid lines) and ISO (2003) absolute thresholds (dotted line). The values below 20 Hz correspond to the ones estimated by Møller and Pedersen(2004).

Harmonic distortion was measured at 1/3-octave intervals using sinusoids with frequencies between 5 and 100 Hz, and at a level of 130 dB SPL. This level is much higher than the levels actually used in the experiment. In all cases, the 2nd, 3rd and 4th harmonics had levels at least 30, 40, and 50 dB, respectively, below that of the fundamental (typically by more than 10 dB). Since the levels used in the experiment were typically at least 30 dB below 130 dB SPL, harmonic distortion levels were much lower than the levels described above. Thus, it seems likely that harmonic distortion in the acoustic signals was not audible. Signals were sent through the ADAT optical outputs of an RME DIGI 96 PC soundcard to an RME ADI-8 D/A converter, using a sample rate of 48 kHz and 24-bit resolution. The 8 loudspeakers were driven by a Crown Studio Reference I (1160 W) power amplifier.

B. Comparative approach

Since the ELCs were obtained binaurally, the left and right DPIMCs were combined to obtain a single representative curve to use in the comparison. This was done using the binaural loudness summation approach proposed by Sivonen and Ellermeier (2006). Basically, at each point in frequency, the powers of the left and right DPIMCs were added to construct a single combined curve. Since no particular frequency point is used as a reference in deriving the DPIMCs, the summation was done after vertical (i.e. level) alignment of the left and right DPIMCs. The chosen vertical shift was the one that minimized the root mean square (rms) error between the curves, considering frequencies between 30 and 125 Hz. This approach was compared to other options such as a simple dB average of the left and right DPIMCs (that does not need the “rms alignment”), the use of the minimum value between left and right DPIMCs (after rms alignment), and the use of a single curve from one ear, i.e either the left or the right DPIMC. Since none gave particlarly better results than the “power summation”, results describe this latter approach. After obtaining the single DPIMCs, these were adjusted vertically to match the ELCs by minimizing the rms deviation between the curves for frequencies between 30 and 125 Hz.

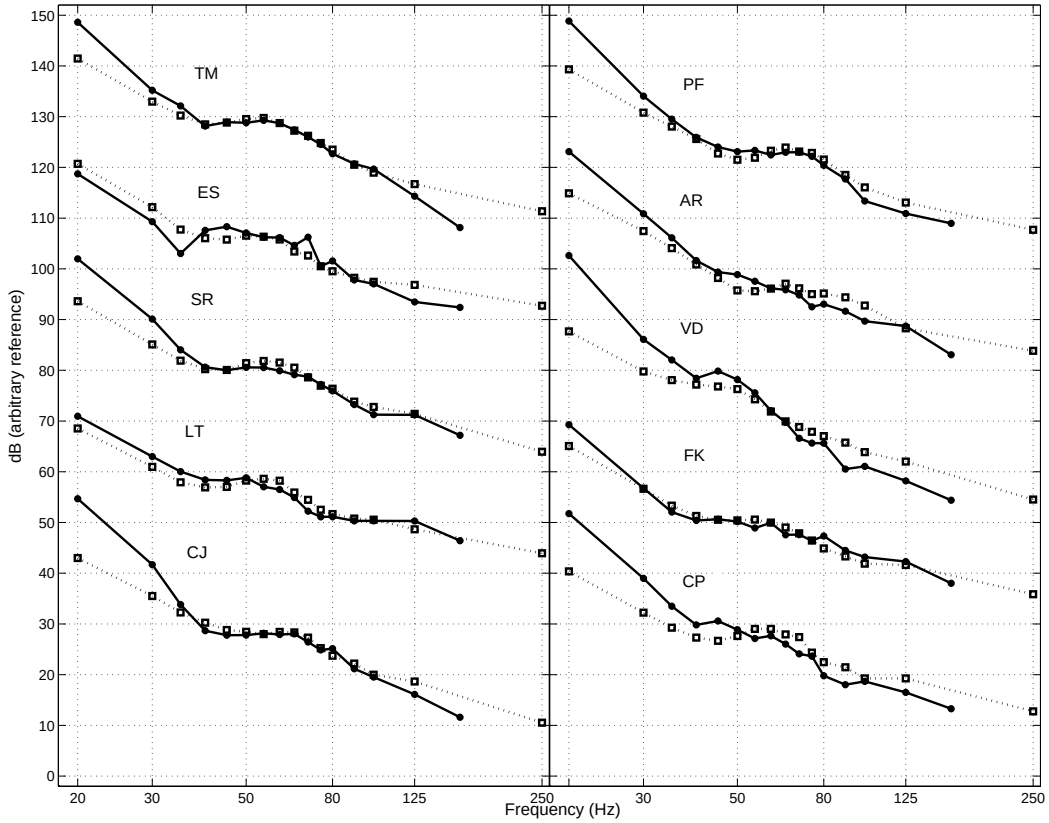

Figure 3. ELCs and fMETFs (the latter combined from left and right curves) after rms alignment (asterisks-solid lines and squares-dashed lines, respectively). The curves are plotted at arbitrary reference.

III. RESULTS

Figures 3 and 4 show the individual ELCs and DPIMCs after the rms alignment (asterisks-solid lines and squares-dotted lines, respectively). Despite evident individual differences, the ELCs appear to have 3 regions: (1) a region above about 100 Hz, shallower in slope; (2) a transition region between about 40 and 100 Hz, more irregular in shape, often resembling a small resonance feature; and (3) a steeper region below about 40 Hz. In what follows, the regions will be termed region 1, 2, and 3, respectively, and the DPIMCs will be directly referred to as fMETFs. The overall features observed in the ELCs have been described previously in Jurado *et al.* (2010a). They agree qualitatively with those of the fMETFs. The latter transfer functions are treated in more detail in another manuscript, where, to account for the data, a lumped element model of the apical cochlea is also presented (Marquardt and Jurado, 2010).

As can be visually inferred, the degree of similarity between both curves varied somewhat, depending on the subject. However, in general, whithin region 2 the curves are close to each other. This was particularly so for subjects TM, SR, LT, CJ, PF, FK, and NL, where the curves have very similar shape and are rougly on top of each other. In region 3, the ELCs are clearly steeper than the fMETFs, although the degree of this difference varied across subjects (see e.g. FK and compare to NL). The mean steepness of the ELCs and

fMETFs in this region was about 21 and 14 dB/oct, respectively. In region 1 it was roughly 11 and 7 dB/oct, respectively. Some individual variation is evident as to where region 3 (the steeper region) started. For example it is roughly close to 45 Hz in AR and PF, while for AD and ES it starts lower in frequency, closer to 30 and 35 Hz, respectively. For some subjects, the transition region (i.e. region 2) in the ELCs took the form of a small resonance feature, such as for TM, ES, SR, LT, PF, VD, and CP. This was in most cases reflected too in the fMETFs, although not so evidently for

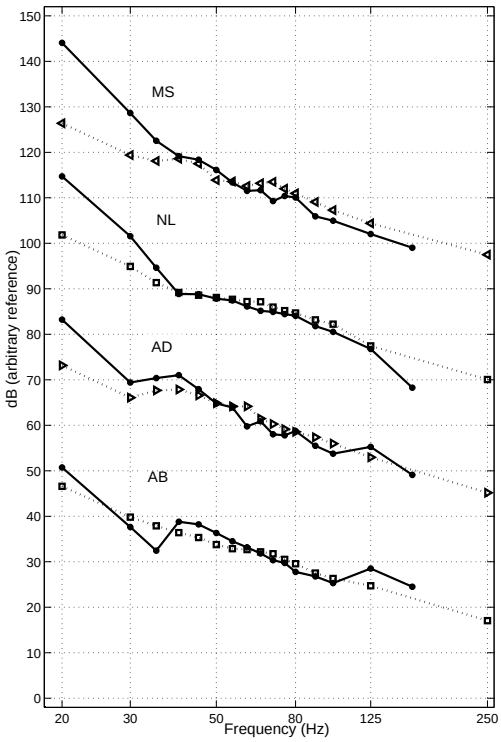

Figure 4. As figure 3 for 4 other subjects. The left and right pointing arrows indicate cases where the fMETFs could be obtained for the left and right ear, respectively.

ES and VD (however VD fits much better with a 6 dB/oct turn, as discussed below). This is noteworthy, since the curves assume a very specific and subject-dependent shape. For other subjects, the transition region in the ELC was smoother or did not assume an evident resonance shape, such as AR, FK, MS, NL, AD, and AB. The relatively smooth transition regions in the fMETFs of e.g. FK and NL were fairly represented in their ELCs. For the cases where a 'hump' feature is evident, the frequency location of this characteristic varied noticeably. For example for TM it covered a range of at least about 40 and 100 Hz, while for PF the hump locati on was higher in frequency, approx from 55 to 100 Hz. As a measure of the goodness of the match, the rms deviation between the curves calculated in the frequency range from 30 to 125 Hz is given in Table I (see 'No turn' column).

Since it became evident that some added steepness in the fMETFs was needed to account for the steepness of the ELCs, particularly in region 3, the fMETFs were turned by 6 dB/oct below 40 Hz, and a new comparison was done. The turn may account for the probable effect of inner-hair-cell response patterns (Cheatham and Dallos, 2001). This is further discussed in section IV.

Table I: Root mean squared (RMS) deviation between ELCs and fMETFs. 'No turn' consider the original fMETFs, and 'Turned' indicate that a 6 dB/otc turn was applied to the fMETFs.

Subject	RMS Error (dB)	
	No turn	Turned
TM	1.1	0.7
ES	2.2	2.7
SR	1.7	1.0
LT	1.5	1.2
CJ	2.0	1.4
PF	1.6	1.2
AR	2.1	1.8
VD	3.1	1.0
FK	1.1	1.4
CP	3.1	2.6
MS	3.3	1.8
NL	2.2	1.5
AD	2.2	2.0
AB	2.4	2.8

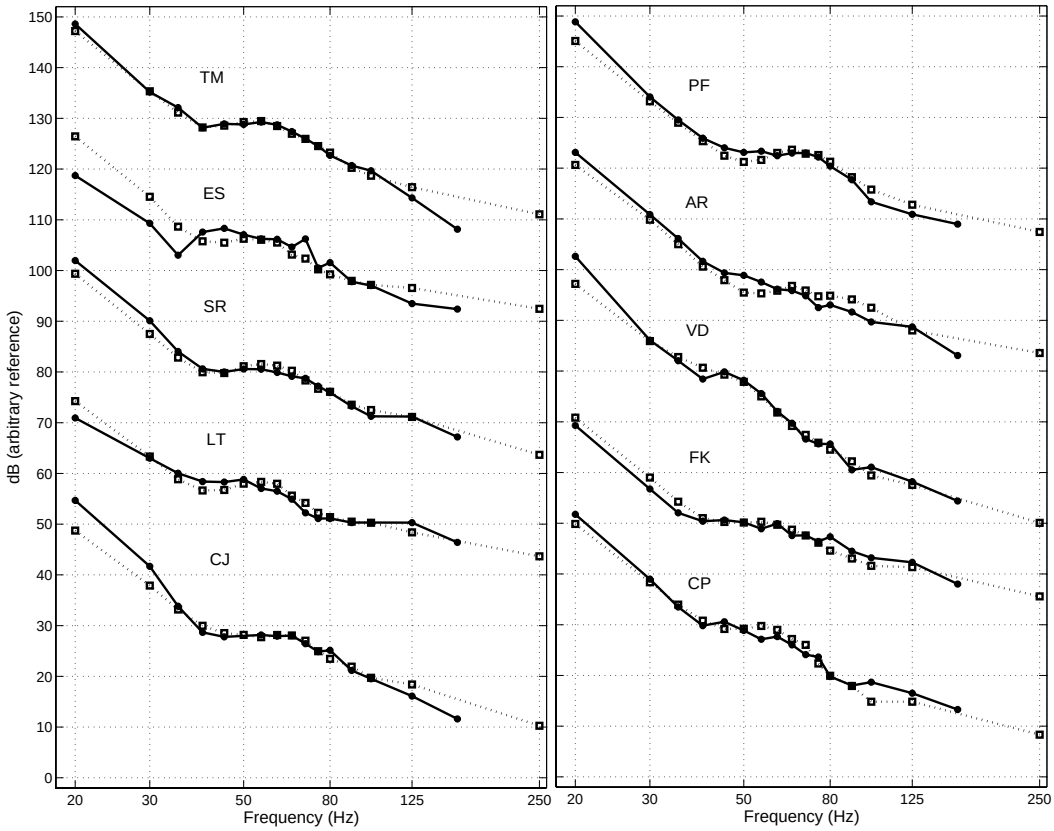

Figure 5. As Figures 3 and 4, but considering the 6 dB/oct turn to the fMETFs.

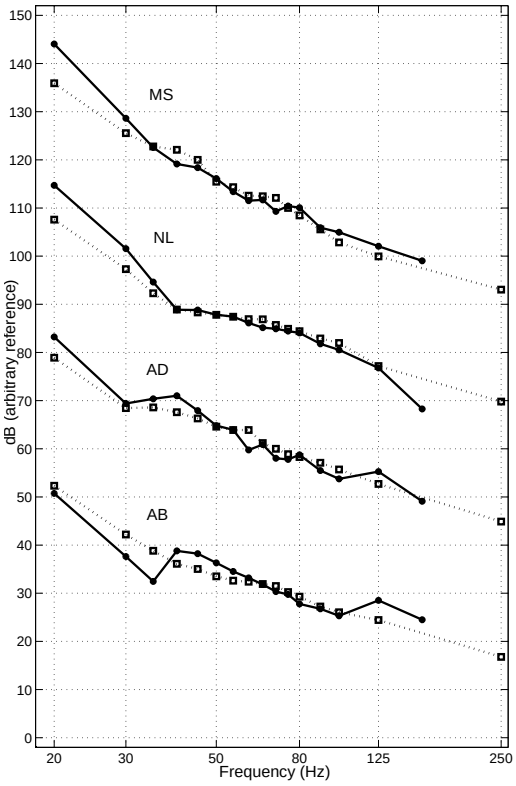

Figure 6. As figure 5, for 4 other subjects.

Figures 5 and 6 show the individual ELCs with the fMETFs turned by 6 dB/oct. The turn was applied down from below 40 Hz in all cases but VD, CP, and MS. For the latter cases, it was noticed that a turn above in frequency, at 100 Hz, represented the data much better. As the figures show, the turn in the fMETFs produced a fairly accurate overall match to the perceptual ELC data. The slopes below about 40 Hz, although still generally slightly steeper in the ELCs, are better accounted for in this manner. For cases where the turn was applied down from 100 Hz, such as VD and CP, the match was much improved within the transition region too (i.e. in region 2). The rms deviation between 30 and 125 Hz is also given for each of these cases in Table I ('Turned' column). In most cases smaller deviations were observed by applying the 6 dB/oct turn (see e.g. VD and MS).

IV. DISCUSSION

In 11 out of 14 cases, the rms deviation between the ELCs and fMETFs was ≤ 2 dB (considering the 6 dB/oct turn to the fMETFs). This is a relatively small quantity and is considered a very satisfactory match, even more so if one considers inherent limitations in the method. For example, being a psychophysical measure, the ELCs can be affected by factors difficult to control, such as psychological changes in criteria and inherent difficulties in comparing signals with different frequencies. The quality of the signals appeared indeed to vary considerably, especially when approaching the infrasound region.

Besides, another factor that may introduce uncertainties is related to the low-frequency cabin. It is an effective pressure field across its whole volume up to about 61 Hz, and as shown in figure 1, some spatial differences in SPL within the listening space are expected at higher frequencies. Therefore, no strong focus has been put into results describing the region above about 125 Hz (i.e. region 1). Factors such as differences in subject's heights relative to the equalization points, or if a subject moves inside the cabin, may affect to some limited extent the result. Furthermore, since the ELCs and fMETFs were measured in different days, comparisons are based on the assumption that the hearing organ state has not changed, i.e. that it is fully time independent. Altough not great overall changes are expected in such a short period (both measurements took place within up to a month from each other), this cannot be assumed to be completely true and is another factor that may introduce uncertainty.

On the other hand, the measurement of fMETFs is subject to noise, and although measures to minimize these effects were taken, it cannot be completely discarded that some noise artifacts may have affected the resulting curves (see Marquardt and Jurado (2010) for details).

Considering all these factors, it is remarkable that such a distinctively close agreement between the two set of curves was found in the majority of cases. This provides manifest indication that loudness for pure low-frequency tones highly depends on the shape of the individual's fMETF. At the same time, the shape of latter transfer function is influenced by the helicotrema shunt mechanism. Possible physiological differences in cochlear anatomy, such as cochlear length and helicotrema size, can help to explain observed individual differences in the shape of the fMETFs (Marquardt and Jurado, 2010). Therefore, loudness of pure low-frequency tones appears to be highly dependent on cochlear anatomy.

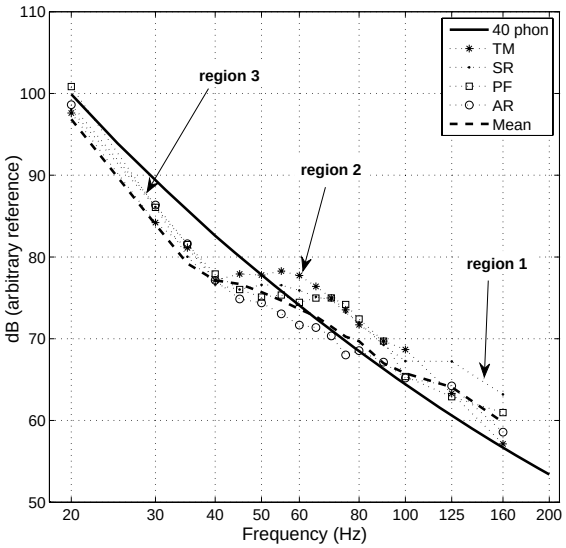

Figure 7. 40-phon ELC and example individual ELCs (solid lines and symbols-dotted lines, respectively). The mean ELC from the 14 subjects is given in dashed lines. Individual and mean ELCs have been vertically displaced for ease in the comparison.

In region 3 (i.e. below about 40 Hz), the ELCs were generally steeper than the

fMETFs. While the ELCs are expected to be affected by inner-hair cell (IHC) response patters (Cheatham and Dallos, 2001), this is not so for the fMETFs, since IHCs are not attached to the tectorial membrane. The 6 dB/oct turn was therefore used to account for the probable dependence of neural spike rates on the velocity of IHCs. This provided a much better fit to the data. The clearly larger steepness of the ELCs (than the fMETFs) below about 40 Hz may help to provide further interpretation of the results of Jurado *et al.* (2010a), who estimated the influence of the fMETF on auditory tuning. They measured psychophysical tuning curves (PTC) for several signal frequencies in the frequency range below 100 Hz. Assuming that the fMETF took a similar shape as the ELCs described here, they estimated its influence on frequency selectivity by excluding its filtering effects from the PTCs. The result of this was that below about 40 Hz the low-frequency skirts of the "treated" tuning curves were practically flat and therefore almost fully determined by the filtering effects of the assumed fMETF. However, had the fMETF been assumed to take the form of the DPIMCs described here, then the low-frequency skirt of the tuning curves would have had a (small) finite slope remaining. Therefore, this suggests that besides the helicotrema shunt mechanism (Jurado and Moore, 2010), there is probably another mechanism (possibly IHC response), that influences frequency selectivity for frequencies below about 40 Hz.

It is relevant to emphasize that in at least half of the cases a resonance feature was clearly observed in region 2 of the DPIMCs and ELCs. This is not present in the ELCs detailed in the ISO 226 (2003) international standard. Figure 7 shows example individual ELCs together with the ISO 226 40 phon ELC (dotted line-symbols and thick-solid line, respectively). The latter curve was chosen merely as example since its values and tendencies were somewhat similar to the measured ELCs. The mean ELC from the 14 subjects is also plotted in thick dashed line for comparison; measured ELCs have been arbitrarily displaced vertically for ease in the comparison. As shown, the ISO curve is smooth (as those of other phon levels) and fails to represent the observed frequency dependence of loudness perception at low frequencies.

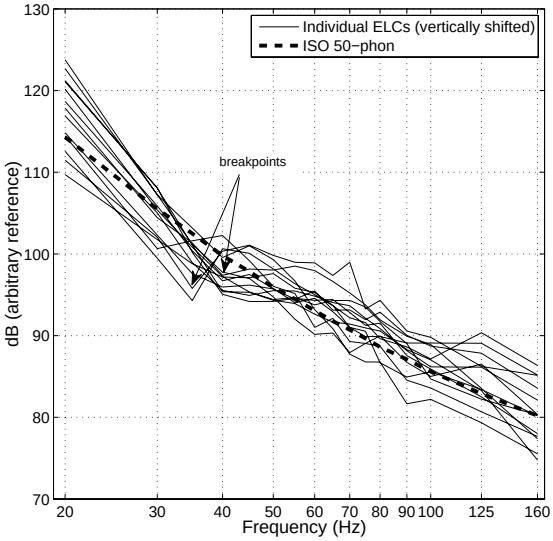

Figure 8. Individual ELCs (solid lines) displaced vertically to fit the 50-phon ELC (thick dashed-line). The breakpoints separating regions 2 and 3 are indicated by the arrows.

Even if region 2 is smooth (see circles, subject FK, or mean ELC, thick-dashed line), the observed increase in slope in region 3 is rather abrupt, and fails to be represented by the standardized curves since their slopes tend to increase smoothly. Discrepancies between the shape and values in individual ELCs and standardized ISO phon curves will be observed on different degrees, mainly in regions 2 and 3, depending on which criteria is used to match the ISO curves. By fitting individual ELCs in their whole frequency range (i.e. 20-160 Hz) to different ISO-phon curves we found differences between the ISO-phon curves

and individual curves were more salient in region 3, and could easily exceed 10 dB. By fitting the curves considering regions 3 and 1 (i.e. the smoother regions in the ELCs, which are more similar to the standardized curves) the best fitting ISO-phon curve was 50 phon, and the largest differences were 9 dB in regions 3 and 2, and 7 dB in region 1.

This latter fit is shown in figure 8. Similarly as shown in figure 7, it can be appreciated that the "breakpoints"between regions 3 and 2 (indicated by the arrows in figure 8) are not represented in the ISO curves; these are also individually different. Furthermore, as described above, region 2 is smooth in the ISO-curves, unlike the shape we observed in the individual ELCs in at least half of the cases.

V. CONCLUSIONS

We have obtained and compared a set of individual ELCs and DPIMCs defined in high resolution in the low-frequency range. These are our main conclusions:

(1) The individual ELCs and DPIMCs present similar overall features which agree qualitatively well. Their frequency dependence can be subdivided into 3 regions: 1) a region above about 100 Hz, shallower in slope; 2) a transition region between about 40 and 100 Hz, more irregular in shape, often resembling a small resonance feature; and 3) a steeper region below about 40 Hz.

(2) In the majority of cases, the shape of the ELCs obtained for the individual subjects matched quantitatively well too with the shape of their DPIMCs, particularly within region 2 (i.e. the transition region). This suggests that loudness perception of low-frequency sinusoids is highly influenced by cochlear anatomy.

(3) Below about 40 Hz, the ELCs were on average about 7 dB/oct steeper than the DPIMCs.

(4) Individual ELCs and the mean ELC for the population considered in this study fail to be adequately represented by the ISO-226 (2003) standardized isophon-curves, since the latter curves assume a smooth variation of loudness with frequency, unlike the trends suggested by the overall evidence presented in this study.

REFERENCES

Cheatham, M. A., and Dallos, P. (**2001**). "Inner hair cell response patterns: implications for low-frequency hearing," J. Acoust. Soc. Am. **110**, 2034-2044.

Frost, G. P. (**1987**). "An investigation into the microstructure of the low frequency auditory threshold and of the loudness function in the near threshold region," J. Low Freq. Noise and Vib. **6**, 34-39.

ISO 226 (**2003**). *Acoustics - normal equal-loudness contours* (International Organization for Standardization, Geneva).

Jurado, C. A., and Moore, B. C. J. (**2010**). "Frequency selectivity for frequencies below 100 Hz: comparisons with mid-frequencies," J. Acoust. Soc. Am. (in press).

Jurado, C. A., Moore, B. C. J., and Pedersen, C. S. (**2010a**). "Psychophysical tuning curves for frequencies below 100 Hz," J. Acoust. Soc. Am. (submitted).

Jurado, C. A., Pedersen, C. S., and Marquardt, T. (**2010b**). "Frequency selectivity at very low centre frequencies: the influence of the helicotrema on individual differences in low-frequency sound perception," in *Proceedings of 14th International Conference on Low Frequency Noise and Vibration and its Control* edited by F. Christensen, and C. S. Pedersen (Aalborg, Denmark), pp. 129-145.

Leventhall, H. G. (**2004**). "Low frequency noise and annoyance," Noise Health **6**, 59-72.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467-477.

Marquardt, T., and Jurado, C. A. (**2010**). "The effect of the helicotrema on low-frequency hearing: I. Human forward middle ear transfer functions," (in prep).

Marquardt, T., and Pedersen, C. S. (**2010**). "The infuence of the helicotrema on low-frequency hearing," in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York), pp. 25-36.

Marquardt, T., Hensel, J., Mrowinski, D., and Scholz, G. (**2007**). "Low-frequency characteristics of human and guinea pig cochleae," J. Acoust. Soc. Am. **121**, 3628-3638.

Møller, H., and Pedersen, C. S. (**2004**). "Hearing at low and infrasonic frequencies," Noise Health **6**, 37-57.

Sivonen, V. P., and Ellermeier, W. (**2006**). "Directional loudness in an anechoic sound field, head-related transfer functions, and binaural summation," J. Acoust. Soc. Am. **119**, 2965-2980.